

# ESSAYS ON MONETARY MACROECONOMICS

## DISSERTATION

zur Erlangung des akademischen Grades  
doctor rerum politicarum  
(Doktor der Wirtschaftswissenschaft)

eingereicht an der

Wirtschaftswissenschaftlichen Fakultät  
der Humboldt-Universität zu Berlin

von

**M.Sc.-Volksw. Anna Almosova**

Präsidentin der Humboldt–Universität zu Berlin:

Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Wirtschaftswissenschaftlichen Fakultät:

Prof. Dr. Daniel Klapper

Gutachter:

1. Prof. Dr. Michael C. Burda, PhD
2. Prof. Dr. Alexander Meyer-Gohde, PhD

Tag des Kolloquiums: 29.01.2019



## Acknowledgements

I would like to thank my supervisor Michael C. Burda for his advice and encouragement. Through all my PhD years his guidance and support helped me to navigate in the ocean of macroeconomic research. At the same time I always had a lot of freedom to independently determine my research agenda.

I am also grateful to Alexander Meyer-Gohde for his useful comments on this thesis, to Mathias Trabandt for his feedback on my research and his advice on the academic job market and to Lutz Weinke for his critical feedback on my work.

I am very thankful to Christopher Sims for inviting me to Princeton and for all his guidance and support. This thesis has highly benefited from the useful comments of Nobu Kiyotaki, Gianluca Violante, Mark Watson and all the participants of the Macroeconomics and Econometrics seminars at Princeton University.

I am thankful to Hermann Elendner for inspiring me to work on digital currencies, to Andreas Tryphonides for his guidance through the technicalities of the macroeconometrics, to Christoph Menzel, my advisor at the German Ministry for Economic Affairs and Energy, and to all my colleges at Humboldt Universität zu Berlin for helpful comments and insightful conversations. I was fortunate to work in two exciting collaborative research projects with Simon Voigts and with Niek Andresen. I would like to thank them for sharing their knowledge and expertise.

I am also grateful for financial support from the Collaborative Research Center 649 and the Schwerpunktprogramm 1764 of the German Science foundation.

Finally I would not have written this thesis without the constant support of my parents, my sister and my significant other.



# Abstract

This thesis consists of three essays on policy relevant topics in theoretical and applied monetary macroeconomics. It addresses several issues that are faced by central banks: forecasting macroeconomic time series, building a model for policy analysis that captures the most important characteristics of the macroeconomic dynamics, studying novel monetary phenomena in the economy.

The first essay describes how to adopt a data-driven method from the machine learning literature - recurrent neural networks - to forecast inflation. It shows that this methods can deliver more accurate predictions than traditional techniques at up to a 12-month-ahead horizon.

The second essay analyzes different formulations of the monetary policy behavior in a New-Keynesian Dynamic Stochastic General Equilibrium model. It extends the traditional formulation suggested by Taylor (1993) by allowing for asymmetric coefficients in the policy reaction function. Empirical evidence from US data is used to evaluate this extension. According to the results the reaction of the US central bank to inflation is stronger when inflation is above the target and reaction to the output growth rate is stronger when it is below the target.

The last part of this thesis is devoted to the analysis of a new challenge for monetary policy - the emergence of decentralized private digital currencies. The third essay first presents a monetary model of the competition between digital and fiat currencies. It derives the conditions under which the competition from private digital currencies imposes a restriction on monetary policy behavior. Second, the essay develops a monetary model of a private digital currency that circulates according to the blockchain protocol. The model is used to analyze how the blockchain characteristics affect monetary equilibrium in the model.



## Zusammenfassung

Diese Dissertation umfasst drei Essays, die sich mit originellen und politikrelevanten Fragen im Bereich theoretische und empirische monetäre Makroökonomie befassen. Sie analysiert unterschiedliche Bereiche der Arbeit einer Zentralbank: die makroökonomische Prognose, die Analyse der Geldpolitik in einem makroökonomischen Modell, das die wichtigen Eigenschaften der ökonomischen Dynamik erfasst sowie neuere Phänomene im Währungssystem.

Der erste Teil der Dissertation beschreibt, wie Rekurrente Neuronale Netze - eine Datenbasierte Methode aus dem Bereich Maschinelles Lernen - für makroökonomische Prognosen angewendet werden kann. Die Ergebnisse des ersten Essays zeigen, dass diese Methode präzise Vorhersagen der Inflation in 1 bis 12 Monaten liefern kann.

Der zweite Essay analysiert die Formulierung von monetären Regeln in Neu-Keynesianischen dynamischen stochastischen makroökonomischen Gleichgewichtsmodellen. Er untersucht mit Hilfe der Daten eine Erweiterung der Taylor-Regel (Taylor, 1993), die Asymmetrien in der Reaktion der Geldpolitik zulässt. Die Untersuchung zeigt, dass die Reaktion der Zentralbank in den USA tatsächlich asymmetrisch ist. Die Zentralbank ergreift stärkere Maßnahmen, wenn die Inflation höher als die Zielinflation ist, und reagiert weniger, wenn die Inflation niedriger als die Zielinflation ist. Gleicherweise ist die Reaktion der monetären Politik stärker bei zu geringem Produktionswachstum als bei zu hohem.

Der letzte Kapitel der Dissertation befasst sich mit digitalen dezentralen Währungen. Diese privaten Währungen stellen eine neue Herausforderung an Zentralbanken da. Der letzte Essay präsentiert zwei monetäre Modelle, die für eine Analyse der digitalen dezentralen Währungen geeignet sind. Ein Modell beschreibt den Wettbewerb zwischen der Währung der Zentralbank und digitalen Währungen. Der Essay analysiert die Bedingungen, unter denen der Wettbewerb einige Beschränkungen für die Geldpolitik darstellt. Das zweite Modell integriert eine digitale Währung, die auf der Basis vom "Blockchain"-Protokoll zirkuliert. Mit Hilfe des Modells wird untersucht, wie die Eigenschaften der Blockchain auf das monetäre Gleichgewicht wirken.





# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Nonlinear Inflation Forecasting with Recurrent Neural Networks</b>	<b>5</b>
2.1	Introduction . . . . .	5
2.2	Methodology . . . . .	9
2.2.1	Forecasting Models . . . . .	10
2.2.2	Optimization Algorithm . . . . .	12
2.2.3	Data . . . . .	13
2.3	Results . . . . .	14
2.4	Sensitivity Analysis . . . . .	16
2.5	Layer-wise Relevance Propagation . . . . .	19
2.6	Real Time Forecasting . . . . .	21
2.7	Conclusion . . . . .	23
2.8	Appendix . . . . .	27
<b>3</b>	<b>Asymmetries in Interest Rate Dynamics</b>	<b>37</b>
3.1	Introduction . . . . .	37
3.2	Evidence on Asymmetry in Interest Rate Dynamics . . . . .	42
3.2.1	Interest Rate Dynamics . . . . .	42
3.2.2	Linear Monetary Policy Rule . . . . .	45
3.2.3	QAR Analysis . . . . .	46
3.3	The Model . . . . .	48
3.4	Estimation Results . . . . .	53
3.5	Posterior Predictive Checks . . . . .	59
3.6	Conclusion . . . . .	63
3.7	Appendix . . . . .	65
<b>4</b>	<b>New Challenges for Monetary Policy: Digital Currencies and Currency Competition</b>	<b>71</b>
4.1	Introduction . . . . .	71
4.2	A Model of Digital Currency Competition . . . . .	75
4.2.1	Money Demand . . . . .	75

4.2.2	Private Currencies . . . . .	78
4.2.3	Equilibrium with Private Money . . . . .	79
4.2.4	Equilibrium with Public Money . . . . .	82
4.3	A Monetary Model of a Blockchain . . . . .	85
4.3.1	Matching in the Money Market . . . . .	86
4.3.2	Demand for Money and Transactions . . . . .	87
4.3.3	Miners and Money Supply . . . . .	89
4.3.4	Equilibrium Dynamics . . . . .	90
4.4	Conclusion . . . . .	95
4.5	Appendix . . . . .	98

## List of Abbreviations

AR	Autoregressive
CPI	Consumer Price Index
DSGE	Dynamic Stochastic General Equilibrium
GDP	Gross Domestic Product
LOOCV	Leave One Out Cross-Validation
LR	Learning Rate
LSTM	Long Short-Term Memory
LRP	Layer-Wise Relevance Propagation
MAFE	Mean Absolute Forecast Error
NBER	National Bureau of Economic Research
NN	Neural Network
QAR	Quadratic Autoregressive
RMSFE	Root Mean Squared Forest Error
RW	Random Walk
RNN	Recurrent Neural Network
SVAR	Structural Vector Autoregression
VAR	Vector Autoregression



## List of Figures

2.1	Fully-connected Neural Network with One Hidden Layer ( $p=2$ ) . . . . .	11
2.2	Basic Representation of a Recurrent Neural Network ( $p=3$ ) . . . . .	12
2.3	Sensitivity Analysis, Simple NN. . . . .	18
2.4	Sensitivity Analysis, LSTM. . . . .	19
2.5	LRP plots for NN and LSTM . . . . .	20
2.6	In-Sample Real-Time 1-step-ahead Forecasts, %, 2000:01-2018:12 . . . . .	22
2.7	Real-Time Forecast 2007 . . . . .	24
2.8	Real-Time Forecast 2008 . . . . .	24
2.9	Real-Time Forecast 2009 . . . . .	25
2.10	Real-Time Forecast 2010 . . . . .	25
2.11	LSTM Cell Representation . . . . .	27
2.12	Raw CPI Inflation Data . . . . .	28
2.13	Real Time Forecasts, Forecast Step $h=1:12$ , 2000-2005 . . . . .	29
2.14	Real Time Forecasts, Forecast Step $h=1:12$ , 2006-2011 . . . . .	30
2.15	Real Time Forecasts, Forecast Step $h=1:12$ , 2012-2018:07 . . . . .	31
2.16	Layer-wise Relevance Propagation Analysis for Simple NN, Examples (part 1)	32
2.17	Layer-wise Relevance Propagation Analysis for Simple NN, Examples (part 2)	33
2.18	Layer-wise Relevance Propagation Analysis for LSTM, Examples (part 1) . .	34
2.19	Layer-wise Relevance Propagation Analysis for LSTM, Examples (part 2) . .	35
3.1	Monthly Effective Federal Funds Rate, 1960:01-2018:10. Source: FRED . . .	43
3.2	Histogram and Kernal Density Estimation for Changes in the Interest Rate. .	44
3.3	Taylor Rule versus Real Data, US, 1982-2003. Source: Rabanal (2014) . . . .	45
3.4	Histogram and Kernal Density Estimation for Changes in the Interest Rate under a linear Taylor Rule. . . . .	47
3.5	Linex Adjustment Costs . . . . .	49
3.6	Log-Likelihood at Prior Mode for Different Number of Particles . . . . .	54
3.7	Estimated Reaction of the Interest Rate to Inflation and Output Growth . .	56
3.8	Generalized Impulse Response Functions for 1sd Technology Shock . . . . .	58
3.9	Generalized Impulse Response Functions for Monetary Policy Shock . . . . .	59
3.10	QAR Coefficients: Standard Taylor Rule and Asymmetric Taylor Rule . . . .	62
3.11	Recursive Averages . . . . .	69
3.12	Marginal Densities . . . . .	70

4.1	Money Demand as a Function of the Return on Money. . . . .	92
4.2	Monetary Steady State . . . . .	94
4.3	Plot of the Equilibrium Dynamics . . . . .	99

## List of Tables

2.1	CPI Forecast RMSFE relative to Random Walk . . . . .	15
2.2	CPI Forecast RMSFE and Random Initialization Variance . . . . .	15
2.3	Top Best Models of AR with Parameters . . . . .	16
2.4	Top Best Models of NN with Parameters . . . . .	17
2.5	Top Best Models of LSTM with Parameters . . . . .	17
2.6	Summary Statistics for Real-Time Forecasts Errors . . . . .	22
3.1	Summary Statistics of Interest Rate Changes, % . . . . .	43
3.2	Posterior Estimates for DSGE Model Parameters . . . . .	55
3.3	Posterior odds ratios (for different truncation probability) . . . . .	57
3.4	QAR posterior predictive checks . . . . .	61
4.1	Threshold Values for Cost Parameter . . . . .	100





# 1 Introduction

Monetary policy interventions constitute an important part of macroeconomic policy. Efficient monetary policy contributes to price stability, sustainable economic growth, secure and efficient payments and resilience of the monetary and financial system. The three essays in this thesis analyze different challenges that a central bank faces when developing its policy - forecasting macroeconomic time series, building a model for policy analysis that captures the most important characteristics of the macroeconomic dynamics and studying novel monetary phenomena in the economy. The first essay summarizes the results of a co-authored work on applying machine learning methods to macroeconomic forecasting. The second essay identifies the presence of asymmetries in the central bank reaction function by estimating a nonlinear dynamic stochastic general equilibrium model (DSGE). The first two chapters are empirical and use US data for their analyses. The last part of this thesis is devoted to decentralized digital currencies - a novel monetary phenomenon that became a topic for debate in many central banks. The last chapter develops a theoretical framework that can be used to analyze how digital currencies circulate in the economy and what consequences they imply for monetary policy. All chapters address policy relevant questions that have recently become particularly pertinent for central banks through digitalization, collection of large (online) data sources and advancement in computational methods that allow to solve and estimate large scale macroeconomic models nonlinearly. On the methodological side, all essays move beyond traditional linear framework and conduct theoretical or empirical analysis in a nonlinear macroeconomic model.

The recent financial crisis stressed the importance of economic heterogeneity and nonlinearities in the macroeconomic dynamics. In light of the crisis events larger efforts are now devoted to the collection of reliable detailed statistical data that are available without large time lags. Data science methods are becoming widely used at central banks to analyze these large data sets and deliver adequate analysis to support the process of decision-making. Benoît Cœuré, Member of the Executive Board of the ECB, pointed out in his speech at ECB conference in November 2017, that big data methods can help to overcome the shortcomings of the traditional macroeconomic time-series modeling and provide opportunities for novel economic research.<sup>1</sup>

The first essay in my thesis, titled "Nonlinear Inflation Forecasting with Recurrent Neural Networks" takes a step in the direction of using big data methods to support monetary policy

---

<sup>1</sup>Source: <https://www.ecb.europa.eu/press/key/date/2017/html/ecb.sp171124.en.html>

decision-making. It shows how long-term short memory recurrent neural networks (LSTM RNN) can be applied to inflation forecasting. Accurate economic forecasting lies at the core of efficient macroeconomic policy. Predicting future inflation is one of the key problems of a central bank since inflation is a primary target of most central banks in the world. Recurrent neural networks were developed for sequential data and were highly successful in machine learning applications such as text analysis. However, the macroeconomic application of this method is highly novel. LSTM RNN is a nonparametric estimator that can learn nonlinear regularities in the data. A researcher does not have to specify the form of nonlinearity in advance. The results show that recurrent neural networks can be successfully used for nonlinear macroeconomic prediction. They outperform standard methods, such as the linear autoregressive model, the random walk model and fully connected neural networks on US monthly CPI inflation forecasting. At horizons of one to twelve month ahead, the root-mean-squared forecast errors of the recurrent neural network are approximately one-third of the same measure for the other methods. The analysis also shows that in a real-time forecasting set-up recurrent neural networks deliver significantly more accurate predictions as measured by the Diebold-Mariano (2002) test. The chapter also conducts a sensitivity analysis and provides some visualization of what recurrent neural networks learn from the data by applying a novel layer-wise relevance propagation algorithm.

In the aftermath of the financial crisis many central banks operated in an environment of extremely low interest rates. A big debate among policymakers was concentrated around "liftoffs" of the policy rates. The majority of central banks decide to increase the interest rates gradually and under a cautious monitoring of the economic recovery. For example, in its meeting on December 18-19, 2018, the Federal reserve will decide on the next increase of the Federal funds rate, that would be the ninth interest rate increase since 2015. At the same time at the outbreak of a recession central banks normally quickly cut interest rates to stimulate the economy. Such behavior contradicts the argument of the traditional Taylor (1993) rule which postulates that adjustments of the interest rate are a linear symmetric function of the inflation and output deviations from their targets. The second essay in this thesis, "Asymmetries in Interest Rate Dynamics", investigates whether the central banks' reaction is systematically asymmetric. It augments a New-Keynesian DSGE model with a flexible formulation of the Taylor rule that allows for asymmetry in its coefficients. This formulation nests a linear rule as a special case. I solve the model by second-order perturbations and estimate the model's parameters by Bayesian methods with the help of the particle filter. My results indicate that

the asymmetry coefficients in the Taylor rule are significantly different from zero. It means that significant asymmetries are present in the reaction function of the US central bank. Its reaction to inflation is stronger when inflation is higher than the target and reaction to the output growth is stronger when the output growth rate is below the target. Another finding of this chapter is that the asymmetric reaction function of the Federal reserve cannot be the sole reason for the observed asymmetric interest rate dynamics. The large asymmetry in the interest rate data results from a combination of the asymmetric monetary policy rule and the asymmetry stemming from the real side of the economy.

A novel phenomenon of modern monetary systems - digital currencies - are addressed in the last chapter of this thesis, entitled "New Challenges for Monetary Policy: Digital Currencies and Currency Competition". This study is motivated by the recent emergence of more than 2000 of private decentralized currencies that operate according to the "blockchain" protocol. Several central banks expressed their concerns about potential threats and challenges that digital currencies imply for monetary policy. Some central banks, for example, the Bank of England or Sveriges Riksbank, decided to investigate the potential of central bank digital currency (CBDC), a concept that would incorporate advantages of both decentralized digital currencies and the central bank legal tender. Accurate analysis of digital currencies and CBDCs require a monetary model that takes into account the most important features of the creation and operation of digital currencies.

The third essay develops two theoretical monetary models that can be used to study the issues related to digital currencies. The first part of the essay extends the Lagos and Wright (2005) monetary framework to include the costs of digital currency operation - mining costs. The presence of mining costs makes blockchain currencies fundamentally different from fiat money which can be created at near-zero costs. The model is then used to study the outcome of currency competition between private digital currencies and the central bank fiat money. I show that the presence of mining costs affects the model conclusions. Currency competition plays a role for the equilibrium inflation of the central bank money only if the mining costs are sufficiently low.

The second part of the essay analyzes a monetary system with a single blockchain-based digital currency in circulation. I use the search and matching approach to model how miners and transaction requests "meet" on the money market. In contrast to traditional monetary models, the findings of the chapter indicate the presence of a nonlinear hump-shaped dependency between the price of money and money demand. The nonlinearity

emerges from a "precautionary motive" of the buyers who take money market externalities into account- not all transaction requests "meet" with a miner and are therefore processed. As a result of this nonlinearity, the monetary equilibrium in the model is stable for some calibrations. In addition to developing a novel monetary model of the blockchain operation, this part of the third essay shows how the equilibrium inflation rate is determined by mining costs and rewards to miners.

This thesis makes some progress in integrating nonlinear methods into policy-making and developing new monetary models to address novel policy relevant issues. It takes a step towards a richer and state-of-the-art central bank analysis that could contribute to efficient monetary policy design.





## 2 Nonlinear Inflation Forecasting with Recurrent Neural Networks

### Abstract

This paper demonstrates the value of machine learning techniques in forecasting macroeconomic time series. We show that recurrent artificial neural networks trained with adaptive gradient descent (ADAM) outperform linear autoregressive (AR) and random walk (RW) models in forecasting monthly US CPI inflation for horizons of 1 to 12 months ahead. Performance of a simple fully-connected neural network (NN) is on a par with the linear AR model when the latter is trained with regularization and superior to the RW while a long short-term memory (LSTM) recurrent neural network has a higher forecasting power than both the RW and AR models. At all horizons root mean squared forecast error (RMSFE) of the LSTM is approximately one third of the RMSFE of the RW. We also show that rolling-window real-time forecasts of the LSTM are significantly more accurate compared to the AR and NN forecasts. Additionally we conduct a sensitivity analysis with respect to hyper-parameters and provide a qualitative interpretation of what the networks learn by applying a layer-wise relevance propagation technique (LRP).

**Keywords:** Recurrent Neural Networks, LSTM, Forecasting, Inflation

**JEL classification:** C45, C53, E37

*This chapter is based on the joint work with Niek P. Andresen, a Master's student in Computer Science at Technical University of Berlin, Department of Computer Engineering and Microelectronics.*

### 2.1 Introduction

Accurate inflation forecasting is essential for many economic decisions. Private investors need to predict future inflation to adjust their asset holdings, firms need to forecast the aggregate inflation level to adjust their prices and maximize profits, central banks need to forecast inflation to conduct an efficient monetary policy. Fiscal authorities need to forecast inflation dynamics if they use the rate of inflation to adjust social security payments and income tax brackets.

Inflation forecasting is an interesting yet challenging task. As Stock and Watson (2007) point out, inflation in the US has recently become both easier and harder to predict. On the

one hand, since the mid-80s inflation became less volatile and as a result easier to predict. On the other hand, it became harder to outperform a naive univariate random walk-type forecast. For example, Atkeson and Ohanian (2001) show that averaging over the last 12 months gives a more accurate forecast of the 12-month-ahead inflation than a backward looking Phillips curve. The macroeconomic literature replies to this challenge by arguing that the inflation process might be changing over time. Consequently, a nonlinear model would give a more accurate inflation forecast.<sup>2</sup>

This paper evaluates the performance of a nonlinear nonparametric method from the machine learning literature, that is novel in this context - a long short-term memory recurrent neural network (LSTM RNN) which is a special case of a standard recurrent neural network (RNN). We see four main advantages of this method. First, LSTMs are flexible and data-driven. It means that the researcher does not have to specify the exact form of the nonlinearity. Instead the LSTM will infer it from the data itself. Second, as stated by the universal approximation theorem, under some mild regularity conditions neural networks (NNs) of any type, including LSTM, can approximate any continuous function arbitrarily accurately (Cybenko, 1989). At the same time these models are more parsimonious than many other nonlinear time series models (Barron, 1993). Third, LSTM RNNs were developed specifically for the sequential data analysis and were shown to be very successful with this task. In the machine learning literature this method is widely applied in text analysis (Sundermeyer et al., 2012). For example, smartphones use RNNs to predict the last word in the sentence and help the user while she is typing a message. Finally, the recent development of the optimization routines and the libraries that employ computer GPUs<sup>3</sup> made the training of LSTM RNNs significantly more feasible.

We compare the performance of an LSTM with a simple fully-connected NN model, an AR(p) model and a RW-type model on monthly inflation forecasting on the US data. One must note that the performance of all neural networks is determined by several hyperparameters such as the number of hidden units or the share of the data that is assigned as training set. To provide some guidance on how to choose these parameters we conduct an extensive sensitivity analysis. In total we train about 4420 different models. Moreover,

---

<sup>2</sup>For example Stock and Watson (2007) fit an integrated moving average (time-varying trend-cycle) model to the GDP-deflator data, and show that the coefficients in this model changed in the beginning of 70s and then again in the mid 80s. The authors conclude that "... if the inflation process has changed in the past, it could change again."

<sup>3</sup>We use Python TensorFlow on 4 nVidia K20m GPUs.



neural networks are often criticized for their "black box" structure. Since NNs are nonlinear in parameters it is difficult to interpret what they actually learn and how the input affects the output. We attempt to assess the relevant importance of the model inputs for the final prediction by applying a layer-wise relevance propagation algorithm (LRP) of Lapuschkin et al. (2016). The idea of this novel method is to decompose the final predicted value into the sum of the positive and negative values coming from the activated hidden units. The outcomes of the hidden units can in turn be decomposed into the contributions of the input neurons. This allows us to track how the value of a particular input contributes to the final prediction of the network.

According to our results, all the considered models: AR when trained with regularization, NN and LSTM outperform the random walk forecasts at all forecasting horizons. One-step-ahead root mean squared forecast error (RMSFE) of the AR, NN and LSTM is approximately 40% of the corresponding measure for the RW model. At longer horizons, 2 through 15 steps ahead, performance of the AR and NN deteriorates with RMSFEs rising from about 55% to 94% of the random walk RMSFEs. However the LSTM continues to make more accurate predictions. At all horizons its errors are significantly smaller than for the AR and equal to approximately one third of the random walk errors. Our findings thus suggest that the LSTM RNN is an efficient nonlinear model for forecasting inflation especially at longer horizons.

Based on our sensitivity analysis we can conclude that the simple NN performs the best when Bayesian information criterion (BIC) is used for the lag length selection and when the maximum number of lags to select from is large. The performance of the NN is insensitive to the number of hidden units as long as there are more hidden units than the number of lags. The results for the LSTM are rather mixed with regard to the information criteria and the maximum number of lags. The best performing model is the one with a large pre-defined number of lags. The accuracy of the LSTM initially increases with the number of hidden units and then plateaus at around 100. For both models, the results are highly insensitive to the learning rate parameter.

Overall, the outcome of the simple NN is more robust whereas the RMSFE of the LSTM can change significantly with different hyper-parameters. We conclude that it is worth one's while to spend time on the fine-tuning of the LSTM model according to the problem at hand. It is also desirable to use a large amount of hidden units in the LSTM RNN. Of course, it comes at the costs of computation time required to train this network.

Our LRP analysis indicates that both NNs and RNNs mostly pay attention to the most

recent lag and the same lag one year ago when computing their predictions. It seems like neural networks are able to detect the annual seasonality in the data and take this into account. Additionally, on average the periods with sharp kinks in the time series contribute a lot to the final prediction. These periods inform the network whether the current trend in the data is positive or negative.

This paper is not alone in applying deep learning methods to macroeconomic forecasting. Ahmed et al. (2010) and Stock and Watson (1998) compared linear and nonlinear methods for macroeconomic forecasting by averaging their performance over a large number of macro time series. Similar to our results these studies find that simple NNs perform well at short forecasting horizons. However, they use a different optimization algorithm to train the NNs and a different procedure to select the test set. Other examples include Kuan and Liu (1995) who demonstrated the success of NNs and RNNs for the exchange rate forecasting in several countries, Swanson and White (1997b,a) who evaluated the advantage of the NNs with time varying coefficients in a real-time forecasting setup, Kock et al. (2011) who discuss direct versus indirect forecasts with NNs.

Chen et al. (2001), Nakamura (2005) and McAdam and McNelis (2005) discuss the inflation forecasting with neural networks. These studies show that NNs outperform benchmark linear models for shorter horizons based on various performance measures. Elger et al. (2006) shows that for shorter horizons RNNs are comparable with Markov switching autoregressive models and at longer horizons Markov switching models are more accurate. Similarly to our analysis they study an RNN for inflation prediction. However, they apply a different type of RNN, different cross-validation and forecast evaluation procedures. Moreover, their study uses GDP inflation data while we focus on CPI inflation forecasting.

Our paper is different from the above mentioned literature in that, to the best of our knowledge, this is the first work that applies an LSTM RNN to inflation forecasting. We also use a different optimization algorithm to fit the NN and RNN models. Moreover, most of the existing papers decide on a particular architecture of the NN a priori. This study, on the contrary, carefully investigates how our conclusions about the comparison of different methods are affected by the hyper-parameters of the NN and the RNN. Finally, our LRP computation is novel to the macro forecasting literature. The discussion of LRP in the machine learning context can be found in Lapuschkin et al. (2016) and Arras et al. (2017). In a broader sense this paper contributes to the literature on the nonlinear time-series forecasting. Teräsvirta (2006) provide an overview of these methods.

The rest of the paper is organized as follows. Section 2.2 describes our set up and the forecasting models. It also gives a brief data description. Results are discussed in Section 2.3. Section 2.4 lays out the sensitivity analysis and section 2.5 presents the layer-wise relevance propagation analysis. Section 2.6 complements our analysis with the real-time forecasting exercise. Section 2.7 concludes.

## 2.2 Methodology

We rank the forecasting methods based on the RMSFE for out-of-sample forecast on the test set. The share of the test set is 10% of the whole data sample. Test sets are constructed by randomly drawing non overlapping data sequences of the length  $p$ , where  $p$  is the selected lag length. The same procedure is applied to obtain validation sets for the sensitivity analysis and real-time forecasts (Monte Carlo cross validation). Note that since our test data samples are randomly drawn from the entire dataset our cross-validation results converge to the leave-one-out cross-validation (LOOCV). However, we can set the number of replications to be smaller than the sample size as it is required for LOOCV.

We focus on the indirect (rolling forward)  $h$ -step-ahead forecasts. While in theory direct forecasts are more robust to model misspecifications, they are less efficient if the model is correctly specified. Marcellino et al. (2006) showed that in the linear forecasting setup indirect forecasts typically perform better than direct ones. Kock et al. (2011) address the same issue for nonlinear prediction methods. They conclude that iterated and direct forecast often have similar performance and their exact ranking is problem- and data-specific. Direct forecasts also require a separate model for each forecasting horizon and are thus more complex computationally. We focus on the indirect forecasts in this study and leave the extension to direct forecasts for our future research.

We are interest in the conditional forecasts at date  $t - 1$  denoted with hats. We start from the following model:

$$y_t = f(Z_{t-1}, W) + u_t, \text{ with } Z_{t-1} = [y_{t-1}, \dots, y_{t-p}] \quad (2.1)$$

where  $y_t$  is a variable of interest,  $f(\cdot)$  is an unknown, potentially nonlinear function,  $W$  is the matrix of model parameters (weights),  $Z$  collects the lags of  $y$ .

One and two step-ahead forecasts based on the information set  $F_{t-1}$  can be computed as:

$$\hat{y}_{t|t-1} = \mathbb{E}[y_t|F_{t-1}] = f(Z_{t-1}, W) \quad (2.2)$$

$$\begin{aligned} \hat{y}_{t+1|t-1} &= \mathbb{E}[y_{t+1}|F_{t-1}] = \mathbb{E}\left[f\left(\hat{Z}_{t|t-1}, W\right) | F_{t-1}\right] = \\ &= \mathbb{E}\left[f\left(\tilde{f}(Z_{t-1}, W) + u_t, W\right) | F_{t-1}\right] \end{aligned} \quad (2.3)$$

where  $\hat{Z}_{t|t-1} = [\hat{y}_{t|t-1}, y_{t-1}, \dots, y_{t-p+1}] = \tilde{f}(Z_{t-1}, W)$

Ideally one would need to estimate the last expectation term by numerical integration. However, moving beyond 2-step-ahead forecast would require evaluating multiple integrals. Moreover, the assumption about the distribution of the  $u_t$  could matter for the result and it would be hard to justify what this distribution should be. Finally, numerical integration or integration by bootstrap could introduce additional inefficiency.

We overcome this problem by assuming that the error term is zero in all states ( $u_t = 0$ ). It implies that our forecast is not an unbiased conditional mean estimator.<sup>4</sup> In other words, our nonlinear models receive solely the information about the first moment of the 1-step-ahead forecast when they compute the  $h \geq 2$  forecasts. It means that if anything we only harm the performance of the nonlinear estimators. Our results can be seen as the lower bound of potential forecasting performance of the nonlinear methods.

### 2.2.1 Forecasting Models

1. RW: random walk-type forecast is constructed as a simple average over the  $n$  previous periods as in Stock and Watson (2007); Aparicio and Bertolotto (2017).

$$\hat{y}_{t|t-1} = \frac{1}{n} \sum_{i=1}^n y_{t-i} \quad (2.4)$$

We tried  $n=3, 6$  or  $12$  and obtain very similar results in terms of model comparison. The exact random walk forecast, that is  $n=1$ , also gives identical results.

2. AR(p): univariate autoregression model of an order  $p$ .

$$\hat{y}_{t|t-1} = A + BZ_{t-1}, \text{ with } Z_{t-1} = [y_{t-1}, \dots, y_{t-p}] \quad (2.5)$$

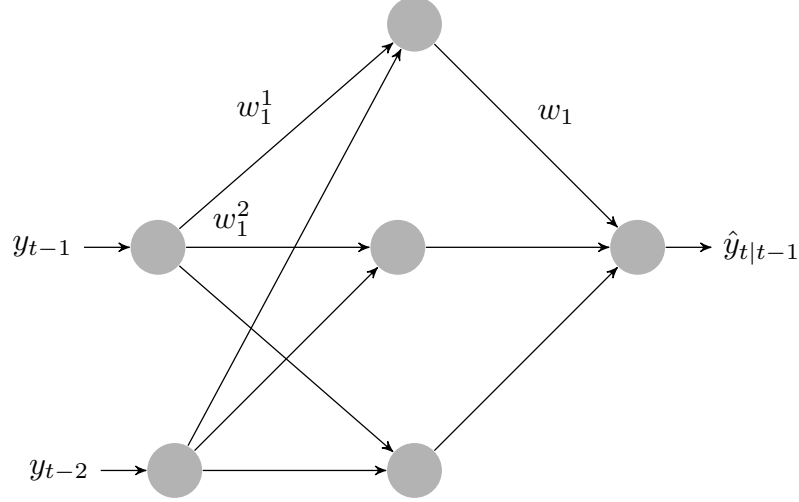
3. NN: simple fully-connected neural network with one hidden layer (Swanson and White, 1997b):

$$\hat{y}_{t|t-1} = b + \sum_{n=1}^N w_n \cdot \sigma(b^n + \sum_{\tau=1}^P w_{\tau}^n y_{t-\tau}) \quad (2.6)$$

---

<sup>4</sup>All the models are fit under the early-stopping rule. The parameters are regularized in order to achieve a better generalization. As a result the estimators are biased by construction in any case.

where  $\sigma(\cdot)$  is a non-linear activation function,<sup>5</sup>  $b$  is a constant or bias of the output and  $b^n$  is a bias of the hidden units  $n$ ,  $w_n$  is a weight from the hidden unit  $n$  to the output,  $w_\tau^n$  is a weight from the lag  $\tau$  to the hidden unit  $n$ . Figure 2.1 sketches the structure of this simple NN model (biases are ignored).



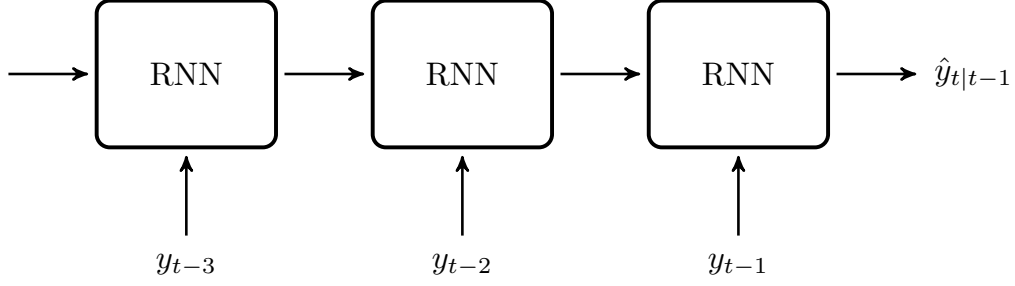
**Figure 2.1:** Fully-connected Neural Network with One Hidden Layer ( $p=2$ )

4. LSTM: Long short-term memory recurrent neural network (Hochreiter and Schmidhuber, 1997).

Representation of a standard RNN is given on the Figure 2.2 where each "RNN" block implicitly contains a simple NN. Basic feature of the RNN is that its prediction is computed sequentially after each time stamp of the input. Intermediate output, which is called the "state" of the network, is used as an additional input at the next time stamp. Recurrent propagation of the state is represented by the horizontal arrows. Note that the weights of the network stay the same, that is, the "RNN" block on the scheme is identical for every time stamp.

RNN's structure has two important implications. First, the network is explicitly informed that the input lag  $y_{t-2}$  comes before the lag  $y_{t-1}$ . More recent lags are likely to be more important for the final prediction. This is in contrast to the simple NN that treats all the lags equally such that the sequence of lags does not matter. Second, the network remembers information about the distant input lags when computing the final output. In text analysis, for example, if an RNN is used to predict the last word in the sentence, the first few words

<sup>5</sup>We use rectified linear unit (ReLU) which is defined as  $\sigma(x) = \max(0, x)$  (Nair and Hinton, 2010).



**Figure 2.2:** Basic Representation of a Recurrent Neural Network (p=3)

can inform the network about whether the sentiment of the sentence is positive or negative. In our application, the state of the RNN can potentially infer the information about trend, cycle or seasonality. Another appealing feature of the RNN which is, however, less relevant in our application, is that the input sequences  $\{y_{t-\tau}\}_{\tau=1}^p$  do not have to be all of the same length e.i.  $p$  can be variable.

The structure of the LSTM network is similar to the RNN except each block has additional "gates". These gates allow the network to decide based on the data what part of the information from the previous step the network wants to use for prediction at the current step. In other words, the state of the network at  $t-1$  is used for the prediction at  $t$  only with some probability. Such architecture leads to a better empirical performance (Hochreiter and Schmidhuber, 1997). A schematic representation of one LSTM block with gates is given in the Appendix 2.8.

### 2.2.2 Optimization Algorithm

All models (except the RW) are trained by backpropagation with an adaptive stochastic gradient descent optimizer - Adam - whose success was largely documented in the machine learning literature (Kingma and Ba, 2014). Adam is different from the standard stochastic gradient descent algorithm in that it updates each parameter  $\theta$  separately and changes the speed of the adjustment  $\eta$  which is called the "learning rate" depending on the "momentum"  $m_t$  or approximated first moment and the "friction"  $v_t$  or approximated second moment of

the gradient  $g_t$ .

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \cdot g_t \quad (2.7)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) \cdot g_t^2 \quad (2.8)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (2.9)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (2.10)$$

$$\theta_{t+1} = \theta_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad (2.11)$$

Here  $\beta_1$  and  $\beta_2$  are tuning parameters of the estimator. Typical values are:  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ .

Parameter gradients are computed at each data point and parameters are then updated accordingly. Once all data points are processed one "epoch" is said to be complete. The number of epochs that are necessary to fit the model varies depending on the problem. Parameter updating is stopped when the RMSFE does not decline anymore for the test set even if it can be further reduced for the training set. This early stopping rule regularizes<sup>6</sup> the parameters and ensures a better generalization<sup>7</sup> of the model. Since the standard frequentist procedure of fitting AR models does not include any regularization we rewrite our AR in a simple NN form and train it in the same way as NN and LSTM models.<sup>8</sup> More specifically, we represent the AR model as a special case of a simple NN with a linear activation function and no hidden layer.

### 2.2.3 Data

We use monthly data on annual US CPI inflation from the FRED database of the Federal Reserve Bank of St. Louis for 1960:01 - 2018:07 which constitutes 703 observations.<sup>9</sup> Inflation rates are calculated as growth rates of the consumer price indices from the same month in the previous year. The data are non-seasonally adjusted. We do not use the seasonally-adjusted series since the seasonality filters can favor nonlinear methods. More specifically, the seasonal adjustment is a non-linear transformation which introduces a nonlinearity in the data by construction. Moreover, seasonality filters are two-sided. Consequently, nonlinear methods can achieve a good fit by simply unraveling the seasonal transformation. Seasonally

---

<sup>6</sup>Regularization essentially means that the method forces the model coefficients to be small when the sample information is weak. See Tikhonov (1963).

<sup>7</sup>The model performance on the data that the model has never seen before.

<sup>8</sup>Alternatively, one can use a Bayesian AR model with a prior that imposes shrinkage.

<sup>9</sup> Future experiments will include: monthly CPI inflation, multivariate forecasts, PCE inflation.

adjusted data are subject to annual revisions and thus cannot be used for our real-time forecasting exercise. We are also interested in testing the hypothesis that neural networks can automatically learn the seasonality of the data.

## 2.3 Results

This section presents the results of the model estimation and compares their forecasting performance on the test sets. We start by accessing the accuracy of our models in terms of the average RMSFE of point forecasts, where the average is computed over 20 runs of Monte-Carlo cross-validation (Dubitzky et al., 2007). To rank our models we proceed as follows. On the first step for each model type we specify a range of possible values for every hyper-parameter. The hyper-parameters include information criteria (AIC, BIC, HQIC or constant pre-defined), maximum lag length to select from (12 or 24), number of hidden units for neural networks (10-100) and the initial learning rate for the Adam optimizer (0.001 to 0.3). In total we train 4420 different models. On the second step we select top 10% of each of the model types: AR, NN and LSTM. For those top candidates we run cross-validation to obtain mean and standard errors of the RMSFE at each forecast horizon. We finally rank the models based on their test errors after the cross-validation and present results for the best performer for each method.

Table 2.1 summarizes the performance for all forecasting models relative to the RW which is taken as a benchmark. The results suggest 3 conclusions. First, at all horizons performance of all models is superior to the simple RW model with forecast errors being on average between 40% and 90% of the RW’s errors. Second, performance of the AR and NN models track each other closely at each forecasting horizon. Since the only difference between the AR and NN is the presence of nonlinear activation function, our finding suggests that regularization leads to a larger accuracy improvements than the nonlinearity of the forecasting model (non-regularized AR model which is not shown here performs very similar to the RW especially beyond the 1-step-ahead forecast). Finally, as the number of forecast steps  $h$  increases, the performance of the AR and NN models deteriorates and their forecast errors become closer in magnitude to the RW. However, the LSTM gives more accurate predictions than all other methods even at the 15-step-ahead horizon. The RMSFE of the LSTM is on average just one third of the RW’s.

Since our networks are trained by a local optimizer multiple optima might be a concern. We follow the literature (Stock and Watson, 1998) and look at the distributions of the forecast



**Table 2.1:** CPI Forecast RMSFE relative to Random Walk

	Test Error	h=2	h=3	h=6	h=12	h=15
AR	0.42 (0.07)	0.57 (0.07)	0.63 (0.08)	0.72 (0.09)	0.86 (0.07)	0.88 (0.07)
LSTM	0.43 (0.06)	0.30 (0.08)	0.26 (0.09)	0.22 (0.08)	0.23 (0.10)	0.27 (0.12)
NN	0.40 (0.05)	0.54 (0.06)	0.64 (0.05)	0.71 (0.05)	0.90 (0.03)	0.94 (0.02)
RW	0.66	0.80	0.93	1.37	2.03	2.23

**Note:** standard errors are computed by 20 runs of Monte-Carlo cross-validation. The values for RW are absolute and given in %.  $h$  indicates the forecast horizon.

errors after training the networks on the same training and test set but with randomly chosen initial conditions for the weights. A low variance of the results after random initialization indicates that the models converge to approximately the same forecast (in terms of RMSFE) in each optimization chain. Table 2.2 presents the absolute values of RMSFE, standard errors (first number in the brackets) and the random initialization variance (second number in the brackets) after this exercise.

**Table 2.2:** CPI Forecast RMSFE and Random Initialization Variance

Model	Test Error	h=2	h=3	h=6	h=12
AR	0.28 (0.05, 0.00)	0.46 (0.06, 0.00)	0.59 (0.08, 0.00)	0.98 (0.12, 0.00)	1.74 (0.15, 0.00)
LSTM	0.28 (0.04, 0.04)	0.24 (0.06, 0.04)	0.24 (0.09, 0.04)	0.30 (0.11, 0.07)	0.47 (0.21, 0.17)
NN	0.27 (0.03, 0.01)	0.43 (0.05, 0.01)	0.60 (0.04, 0.01)	0.98 (0.06, 0.01)	1.82 (0.06, 0.01)

**Note:** Mean and standard deviation of various models after 20 runs of Monte-Carlo cross-validation and after 20 runs with randomized starting values.

AR and NN forecasts have hardly any variance due to the initialization which implies that the training algorithm is robust and finds the same optimum in every run. For the LSTM, on contrary, the multi-start variance constitutes a large portion of the total variation in the forecast. Two conclusions can be drawn from this observation. First, loss function of the best performing LSTM is high-dimensional and highly non-linear. It contains 40901 parameters<sup>10</sup> and is harder to optimize. The outcome of optimization is rather unstable and depends on the stopping points. Nevertheless, even withing these wide bounds the performance of LSTM is unambiguously better at all forecasting horizons.

<sup>10</sup>The best performing LSTM has 100 hidden units which means that at each time step each of its 4 gates transforms 100 states from the previous step and a 1-dimensional input into 100 updated states. Each update also contains a bias parameter. After the last lag of the data the final state of the network is transformed into the final 1-dimensional output by multiplying the state by a (100x1) matrix. A bias is added at this final step. As a result the LSTM has  $4*((101*100)+100)+100+1 = 40901$  parameters.

## 2.4 Sensitivity Analysis

As we discuss in the introduction the performance of different NN models is crucially dependent on the selected hyper-parameters. It can also depend on the lag selection criterion and maximum number of lags that this criterion is allowed to choose from. To access the sensitivity of our results we rank all forecasting models that are used for cross-validation that is the best 10% of each model type and take a closer look at their hyper-parameters.

We find that at 1-step-ahead horizon all the best 10 models are NNs. For the forecasts beyond 2-step-ahead LSTM models are uniformly better. This means that NNs are in general better for short term forecasting and LSTMs for longer horizons and that these results are robust to different hyper-parameter selections. However, the NN's 1-step-ahead RMSFE prediction is very close to the errors of the best LSTM and AR models.

This can also be observed in Table 2.3, Table 2.4 and Table 2.5 which present the top 10 performers among AR models, top 10 among NNs and top 10 among LSTMs respectively. All models are very close in terms of their test errors and the exact ranking is not highly informative. Another observation is that the top 10 AR and top 10 NN models perform very similarly to each other while the performance of the LSTM is affected significantly by the selected hyper-parameters.

**Table 2.3:** Top Best Models of AR with Parameters

	model	Test Error	h=2	h=3	h=6	h=12	h=15	n	inf	p	Lag	LR	epochs
1	AR	0.28	0.46	0.59	0.98	1.74	1.96	20	None	24	24	0.3	16000
2	AR	0.29	0.50	0.67	1.01	1.84	2.08	20	bic	14	24	0.1	16000
3	AR	0.29	0.50	0.66	1.01	1.84	2.08	20	bic	14	24	0.3	16000
4	AR	0.30	0.42	0.54	0.97	1.79	1.98	20	None	24	24	0.1	5000
5	AR	0.30	0.45	0.65	1.17	1.70	1.98	20	aic	19	24	0.1	10000
6	AR	0.30	0.50	0.65	1.00	1.85	2.08	20	bic	14	24	0.3	10000
7	AR	0.30	0.46	0.64	1.16	1.73	1.98	20	aic	19	24	0.3	10000
8	AR	0.30	0.46	0.61	1.00	1.77	1.97	20	None	24	24	0.1	18000
9	AR	0.30	0.49	0.65	0.98	1.84	2.08	20	bic	14	24	0.1	10000
10	AR	0.30	0.51	0.68	1.02	1.85	2.08	20	bic	14	24	0.1	18000

**Note:** selection is based on the average 1-step-ahead RMSFE. h - number of forecast steps, inf - information criterion, p is either the optimally selected number of lags or the max lag, Lag - maximum number of lags, LR is initial learning rate, epochs is the number of optimizer's iterations over the whole data set.

We now document some recommendations regarding the hyper-parameters for each model type. In addition to the tables Figure 2.3 and Figure 2.4 plot the test errors of the NN and LSTM as a function of the number of hidden units, the initial leaning rate and the maximum

**Table 2.4:** Top Best Models of NN with Parameters

	model	Test Error	h=2	h=3	h=6	h=12	h=15	n	inf	p	Lag	LR	epochs
1	NN	0.27	0.43	0.60	0.98	1.82	2.09	50	bic	14	24	0.00	2000
2	NN	0.27	0.42	0.59	0.96	1.81	2.07	100	bic	14	24	0.00	1000
3	NN	0.27	0.42	0.59	0.97	1.81	2.08	100	bic	14	24	0.00	2000
4	NN	0.27	0.43	0.61	0.96	1.83	2.07	50	bic	14	24	0.00	2000
5	NN	0.27	0.44	0.60	0.99	1.84	2.07	20	bic	14	24	0.00	2000
6	NN	0.27	0.43	0.60	0.98	1.83	2.07	100	bic	14	24	0.00	2000
7	NN	0.28	0.44	0.61	0.99	1.84	2.08	100	bic	14	24	0.00	1000
8	NN	0.28	0.42	0.58	0.94	1.78	2.10	100	bic	14	24	0.01	2000
9	NN	0.28	0.44	0.60	0.95	1.82	2.10	50	bic	14	24	0.00	1000
10	NN	0.28	0.45	0.61	0.95	1.81	2.07	50	bic	14	24	0.01	2000

**Note:** h - number of forecast steps, n - number of hidden units, infc - information criterion, p is either the optimally selected number of lags or the max lag, L - maximum number of lags, LR - learning rate, epochs is the number of optimizer's iterations over the whole data set.

**Table 2.5:** Top Best Models of LSTM with Parameters

	model	Test Error	h=2	h=3	h=6	h=12	h=15	n	inf	p	Lag	LR	epochs
1	LSTM	0.28	0.24	0.24	0.30	0.47	0.60	100	None	24	24	0.03	2000
2	LSTM	0.28	0.20	0.19	0.23	0.41	0.51	100	None	24	24	0.00	2000
3	LSTM	0.29	0.24	0.23	0.27	0.43	0.52	100	None	24	24	0.01	2000
4	LSTM	0.29	0.39	0.45	0.57	1.19	1.48	100	aic	19	24	0.01	1000
5	LSTM	0.29	0.38	0.46	0.70	1.38	1.67	100	bic	14	24	0.03	2000
6	LSTM	0.29	0.38	0.45	0.55	1.16	1.42	50	aic	19	24	0.03	2000
7	LSTM	0.29	0.37	0.44	0.55	1.11	1.41	50	aic	19	24	0.01	2000
8	LSTM	0.30	0.40	0.47	0.60	1.18	1.48	100	aic	19	24	0.03	1000
9	LSTM	0.30	0.22	0.23	0.26	0.42	0.55	100	None	24	24	0.10	2000
10	LSTM	0.30	0.38	0.45	0.52	1.09	1.35	50	aic	19	24	0.00	2000

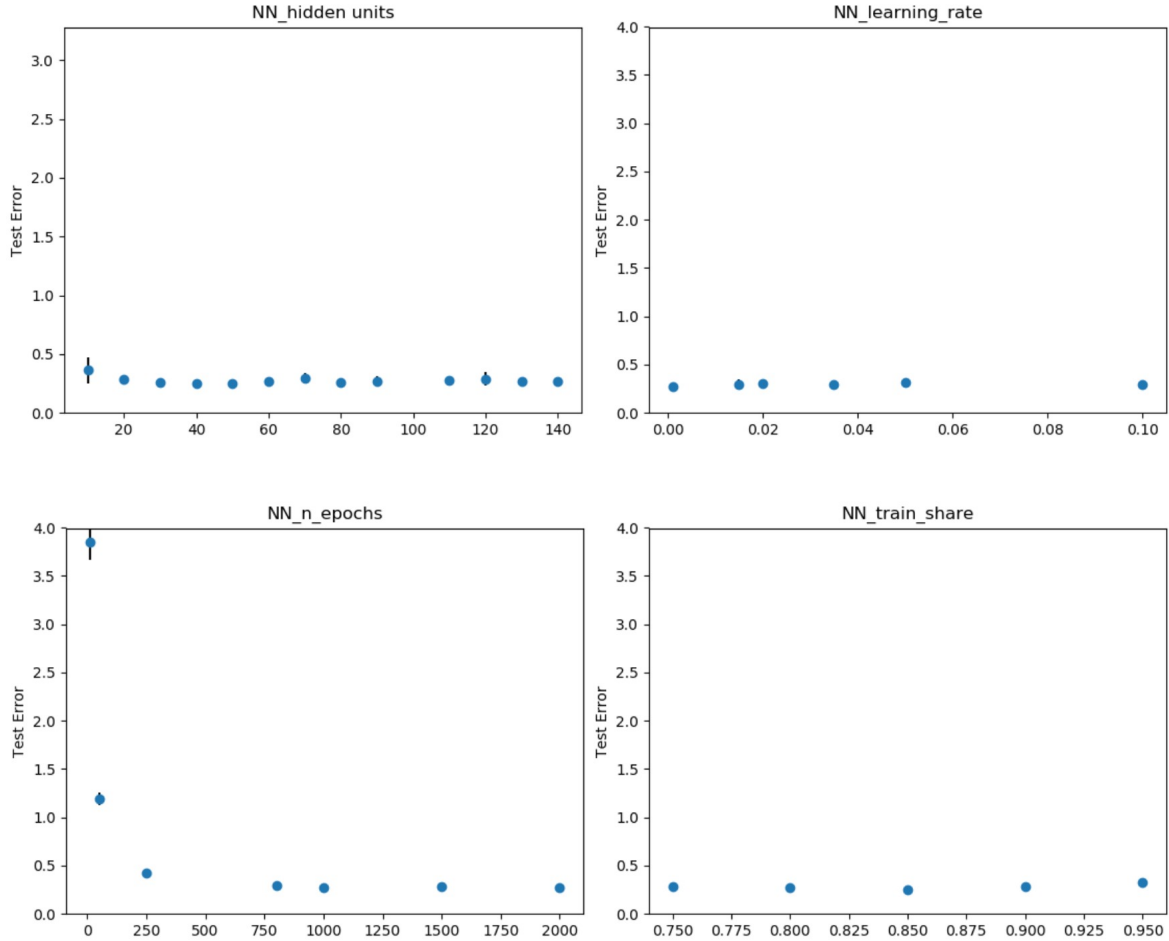
**Note:** h - number of forecast steps, n - number of hidden units, infc - information criterion, p is either the optimally selected number of lags or the max lag, L - maximum number of lags, LR - learning rate, epochs is the number of optimizer's iterations over the whole data set.

number of epochs for which the network can be trained.

*AR model.* AR performs equally well with different the lag selection criterion. Based on the results in Table 2.3 AIC, BIC or fixed number of lags result in virtually the same performance. None of the models with HQIC enters the top 10 list which implies that this criterion does not lead to a good model performance. Learning rate for AR should be at least 0.1 and the network should be trained for at least 5000 epochs.

*NN model.* As can be seen in Table 2.4 the simple NN performs best when Bayesian

information criterion is used for the lag length selection and when the maximum number of lags to select from is large. The performance of the NN is insensitive to the number of hidden units as long as there are more hidden units than lags. Initial learning rate is also irrelevant for the the NN's performance. The bottom left plot on the Figure 2.3 indicates that the test error of the NN achieves minimum at around 1000 epochs and stays unchanged afterwards.<sup>11</sup>

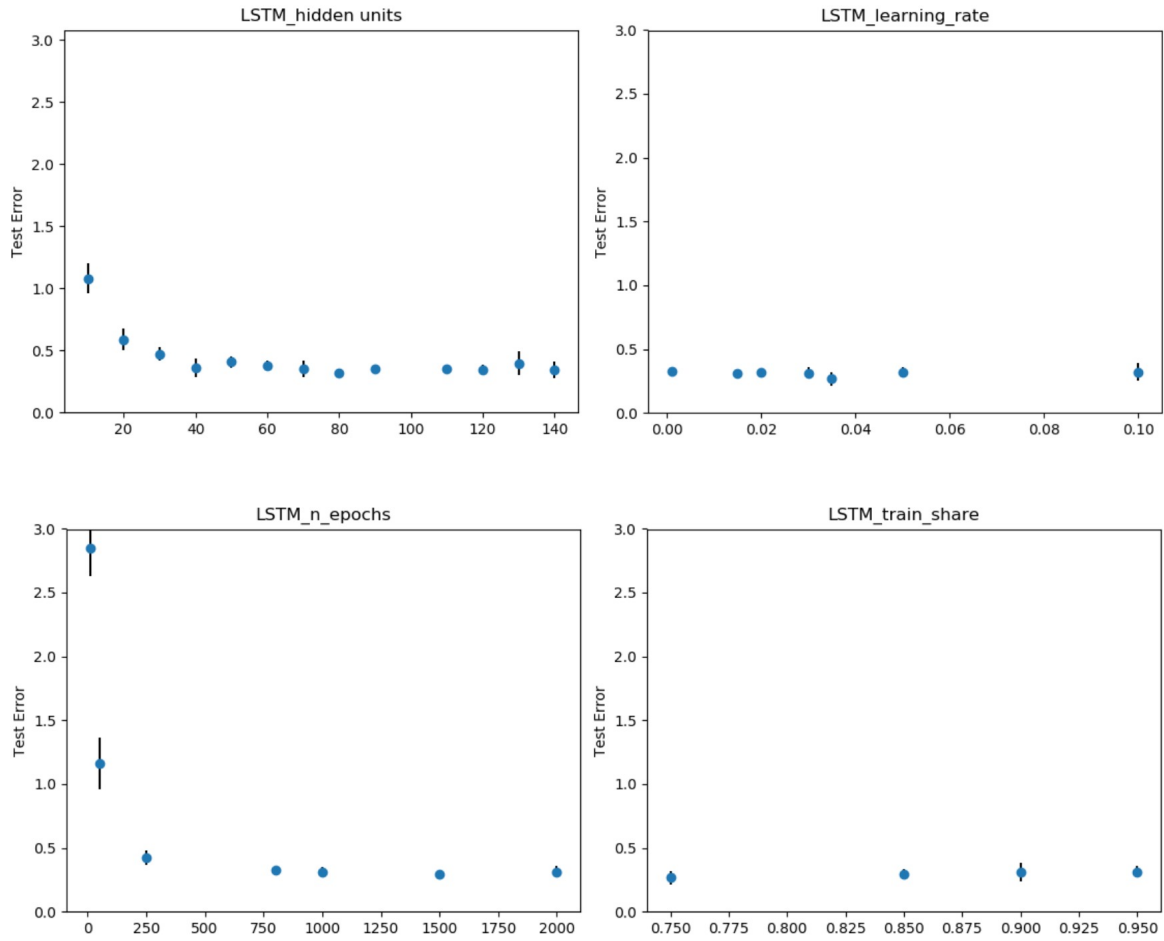


**Figure 2.3:** Sensitivity Analysis, Simple NN.

*Note:* Dots indicate mean RMSFE and black bars indicate 90% credible sets after 20 runs of Monte-Carlo cross-validation.

*LSTM model.* The results for the LSTM are rather mixed with regard to the information criteria. Three best performing models are the ones with a fixed and large number of lags. The maximum number of lags of the top models is 24 which was the largest number we

<sup>11</sup>Note that the train error is strictly decreasing in the number of training epochs while the test error is not - the problem called over-fitting. Training of the networks is therefore stopped at the point of the test error minimum and before the train error arrives its minimum. This represents the "early-stopping" principle.



**Figure 2.4:** Sensitivity Analysis, LSTM.

*Note:* Dots indicate mean RMSFE and black bars indicate 90% credible sets after 20 runs of Monte-Carlo cross-validation.

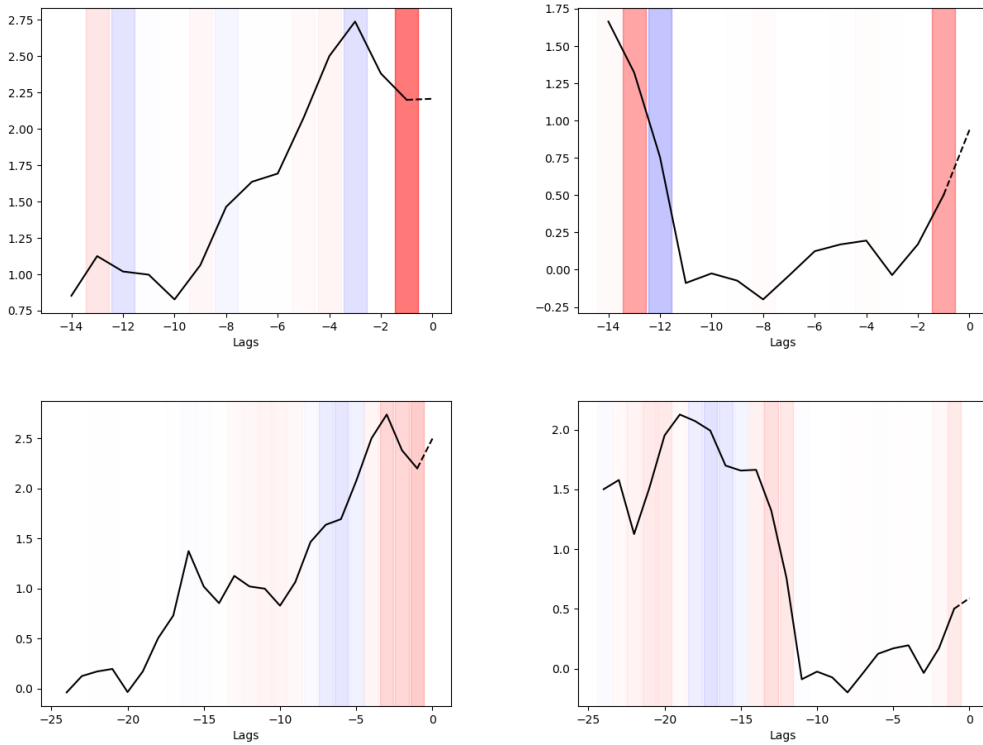
try. The forecast error of LSTM initially decreases with the number of hidden units and then plateaus at around 100. As for the NN the LSTM results are highly insensitive to the learning rate parameter and for both models it is essential to allow for a reasonable number of the training epochs. Best LSTMs are normally trained for 2000 epochs.

## 2.5 Layer-wise Relevance Propagation

After fitting NN and LSTM models one might be naturally interested in interpreting what the networks learned from the data and understanding what features of the input (in our applications different lags) are the most important for the network outcome. In contrast to linear models it is not possible to directly interpret the weights of the neural networks since the final prediction is a nonlinear function of the network parameters.

Lapuschkin et al. (2016) suggest to access the importance of model inputs by layer-wise relevance propagation (LRP). A detailed description of the LRP algorithm lies beyond the scope of this paper and we will only sketch the main idea. The LRP procedure attaches a value to every neuron (including the input neurons - lags) that quantifies its contribution to the network output. This value is called "relevance". It is computed layer-wise by aggregating the signals that a neuron contributes to its successors. The signal's value depends on the weights attached to the connections between a given neuron and its successors. Relevances of the input "neurons" give an estimation of their relative importance for the final prediction.

Figure 2.5 depicts two examples of the LRP measurement for the top performing NN (top row) and the top performing LSTM (bottom row). Additional LRP plots can be found in the Appendix 2.8. The black solid line depicts the actual data. Each bar represents the relevance of a particular lag which was fed into the model as an input. The LSTM has 14 lags and the NN has 24. The rightmost value of each plot is the 1-step-ahead prediction of the corresponding network. Red bars indicate the lags which contribute positively to the predicted values and blue bars imply that the value at the corresponding lag tells the network to decrease the final prediction. Intensity of the color measures the magnitude of the relevance so white bars represent zero relevance.



**Figure 2.5:** LRP plots for NN (top row) and LSTM (bottom row)

Several observations are worth pointing out. The plots for the NN suggest that the most recent lag ( $t=-1$ ) as well as the 12th and 13th lags always contribute the most to the final network outcome. Additionally, the lags with the kinks in the data are sometimes important as, for example, the 3rd lag on the left upper plot. In a nutshell, when computing its 1-step-ahead prediction the simple NN takes the average over the most recent lag and the same lag one year ago and then adjusts the forecasted value if some extreme turns in the trend were observed. It seems like the NN is able to understand the annual nature of the data.

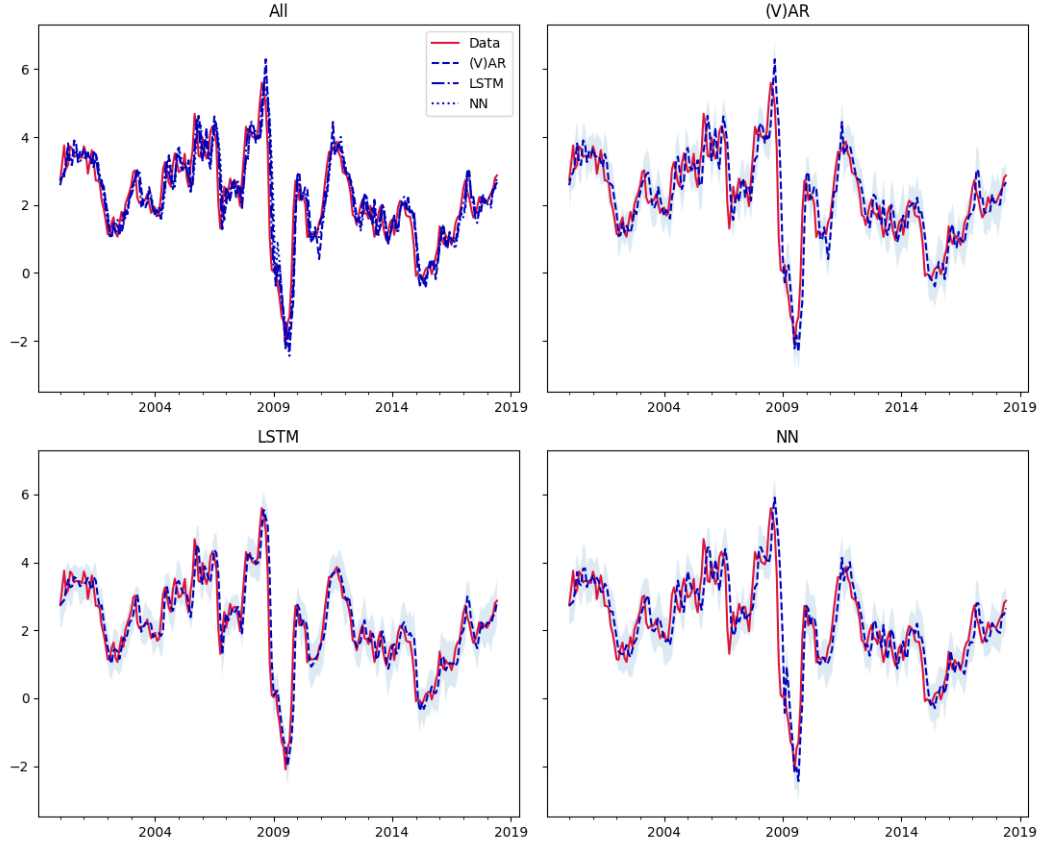
The results for LSTM are more complex. In general the network learns a periodic pattern in which a series of 3-4 blue bars always follow after a series of 3-4 red bars. That is the last four lags always have a positive contribution, 7th through 9th only negative and so on. The magnitude of the negative and positive contributions depends on the inflation level at the corresponding lags. The higher the inflation was 1 month ago the more relevant is this observation for the LSTM prediction. While it is not possible to find a simple rule for how the LSTM decides on its forecast value, our LRP analysis suggests that this recurrent neural network fits a nonlinear periodic function and constructs its prediction based on the presence of peaks at particular lags. Additional plots in the appendix also support these conclusions.

## 2.6 Real Time Forecasting

After documenting a good performance of the LSTM models we conduct a counter-factual exercise to answer the question what would have happened if forecasters were using NN and LSTM models instead of AR(p) to predict inflation since 2000. We recursively compute inflation forecasts year by year. Estimation for the year 2001 are based solely on the information up until 2000, including lag length selection, model fitting and cross-validation to select the hyper-parameters and to compute the confidence bounds. Forecast for  $h \geq 2$  are computed by iterating forward the 1-step-ahead forecast.

*One-step-ahead rolling window forecasts.* On Figure 2.6 we plot the realization of CPI inflation (red solid line) together with the corresponding forecasts (blue dashed lines). Each point on the forecast line is computed recursively as a one-step-ahead point forecast by training the model on the preceding 480 months (rolling window procedure). The series are aligned such that the vertical distance represents the forecast error. Forecast lines are in general similar but do differ in some periods. Visual inspection suggests that the LSTM forecast is on average more accurate. The statistics in Table 2.6 confirm more formally that over the entire forecasting period the LSTM's 1-month-ahead forecast is more accurate both

in terms of MSFE and MAFE (mean absolute forecast error).



**Figure 2.6:** In-Sample Real-Time 1-step-ahead Forecasts, %, 2000:01-2018:12

**Table 2.6:** Summary Statistics for Real-Time Forecasts Errors

	MAFE	MSFE	DM test (AR)		DM test (NN)	
			statistics	p-value	statistics	p-value
AR	0.74	0.75	n.a.	n.a.	n.a.	n.a.
LSTM	0.63	0.56	-5.14	0.00	-3.72	0.00
NN	0.68	0.65	-3.61	0.00	n.a.	n.a.

**Note:** MAFE stands for mean squared absolute error. The comparison model for Diebold-Mariano test is indicated in brackets.

To assess the significance of the sample forecast superiority we conduct the Diebold-Mariano test (2002) with modification suggested by Harvey et. al (1997) to account for serial correlation.<sup>12</sup> Given two forecast series the test computes a forecast error loss differential  $d_{12t} = L(e_{1t}) - L(e_{2t})$  for each  $t$  and runs a  $z$ -test of the hypothesis that the mean of

<sup>12</sup>The code is taken from John Tsang: <https://github.com/johntwk/Diebold-Mariano-Test>.



the differential is zero. The error loss function  $L(\cdot)$  is specified by the researcher; we chose quadratic loss, e.i.  $L(e_t) = e_t^2$ . Under the null, the Diebold-Mariano statistics it is distributed according to a normal distribution:

$$\frac{\bar{d}_{12}}{\hat{\sigma}_{\bar{d}}} \sim N(0, 1) \quad (2.12)$$

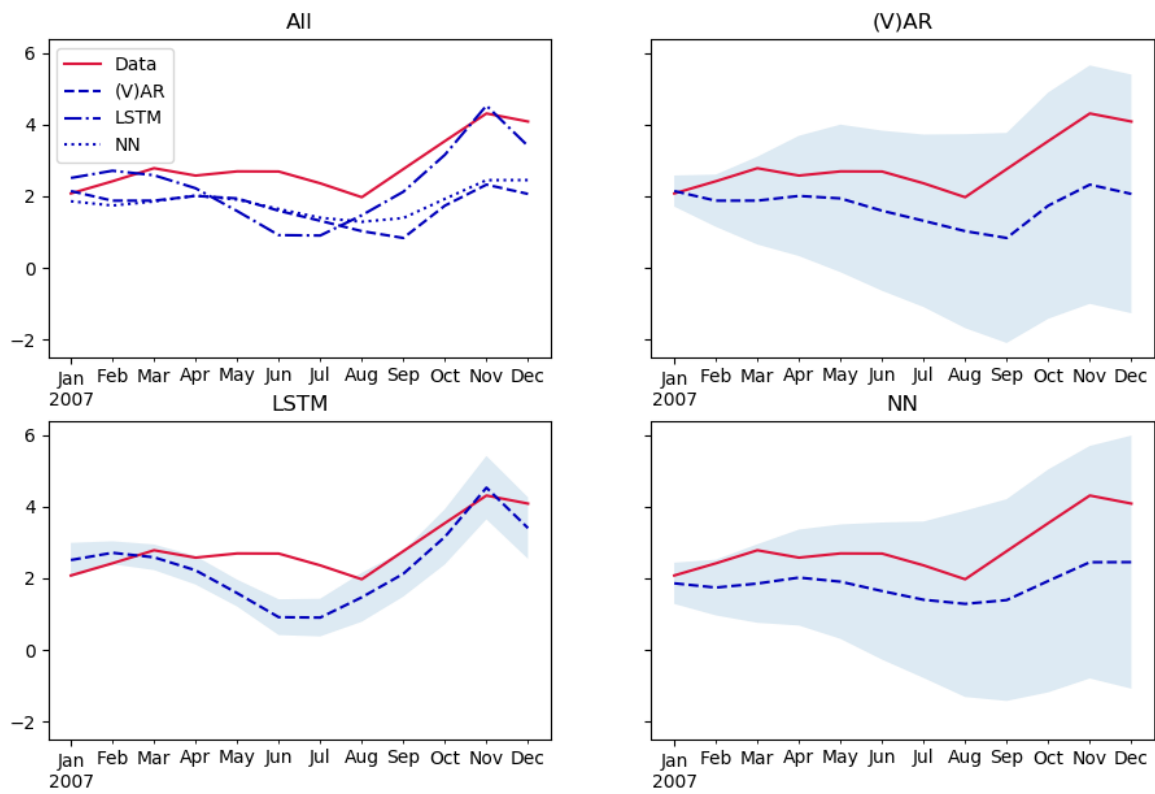
where  $\bar{d}_{12}$  is the sample average and  $\hat{\sigma}_{\bar{d}}$  is the sample variance of the statistics.

Test statistics and p-values are presented in Table 2.6. The results indicate a significant superiority of the LSTM forecasts. Both mean differentials for the LSTM versus NN and versus AR are negative and significant.

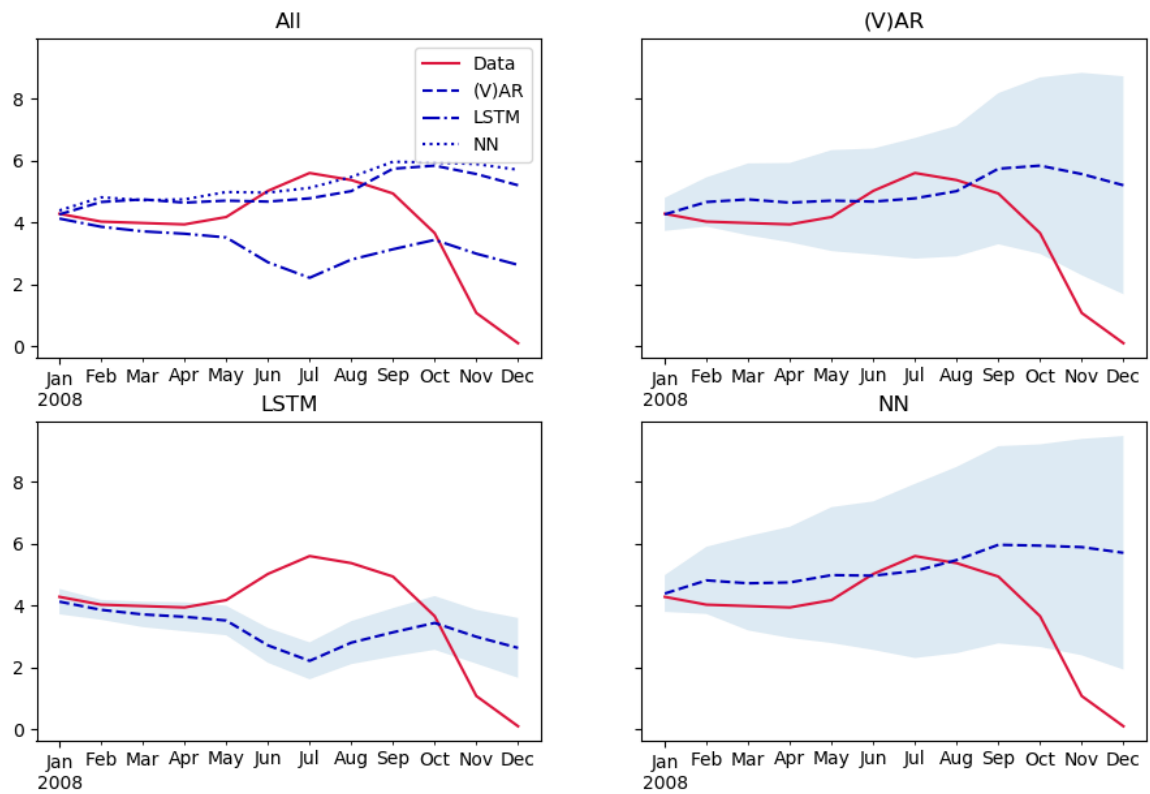
*Longer horizons indirect forecasts.* Figure 2.7 through 2.10 depict several examples of the 1 to 12 month ahead predictions of the different models. Red solid lines depict the real data series and the blue dashed lines are the forecasts. Pictures for all years 2000 through 2018 are presented in Appendix 2.8. These pictures provide some intuition why LSTM predictions for longer horizons are more accurate than the linear forecasts. The LSTM representation of a time series is more nonlinear. This allows the LSTM to better fit unanticipated declines (as for example for the years 2009, 2010). It can also worsen the forecast if there was no trend in the real data (as for example in 2010). NNs produce nonlinear forecasts which are somewhat closer to a straight line. It makes these forecasts more accurate for shorter horizons. However, forecasts can become highly inaccurate at longer horizons. The errors of indirect LSTM forecasts at longer horizons exhibit high variation. LSTMs are extremely accurate in some years, for example in 2007, and produce "wild forecasts" in others, as for instance in 2010. Implausible forecasts can presumably be eliminated by the researcher through forecast trimming (Stock and Watson, 1999).

## 2.7 Conclusion

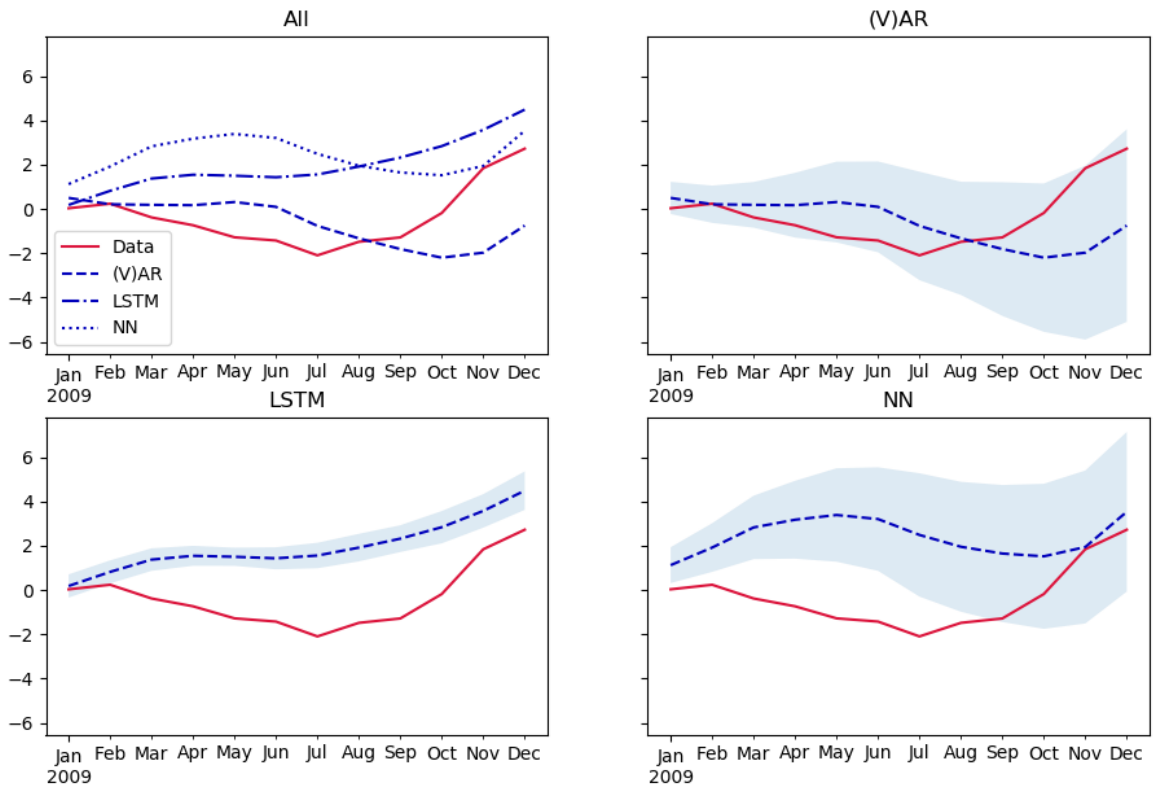
This paper evaluates the performance of a nonlinear machine learning method - long short-term memory recurrent neural network (LSTM) - on the inflation forecasting and compares it to the standard fully connected neural network as well as to the classical methods - the random walk model and linear autoregression models. According to our findings the LSTM outperforms all other forecasting techniques especially at longer horizons with the RMSFE being approximately one-third of the errors for the classical models. We show that LSTM models are sensitive to the choice of hyper-parameters and it is therefore worth one's while to tailor these parameters according to the specific forecasting task. Our LRP analysis suggests some intuition for the reasons behind good performance of the LSTM networks: in addition



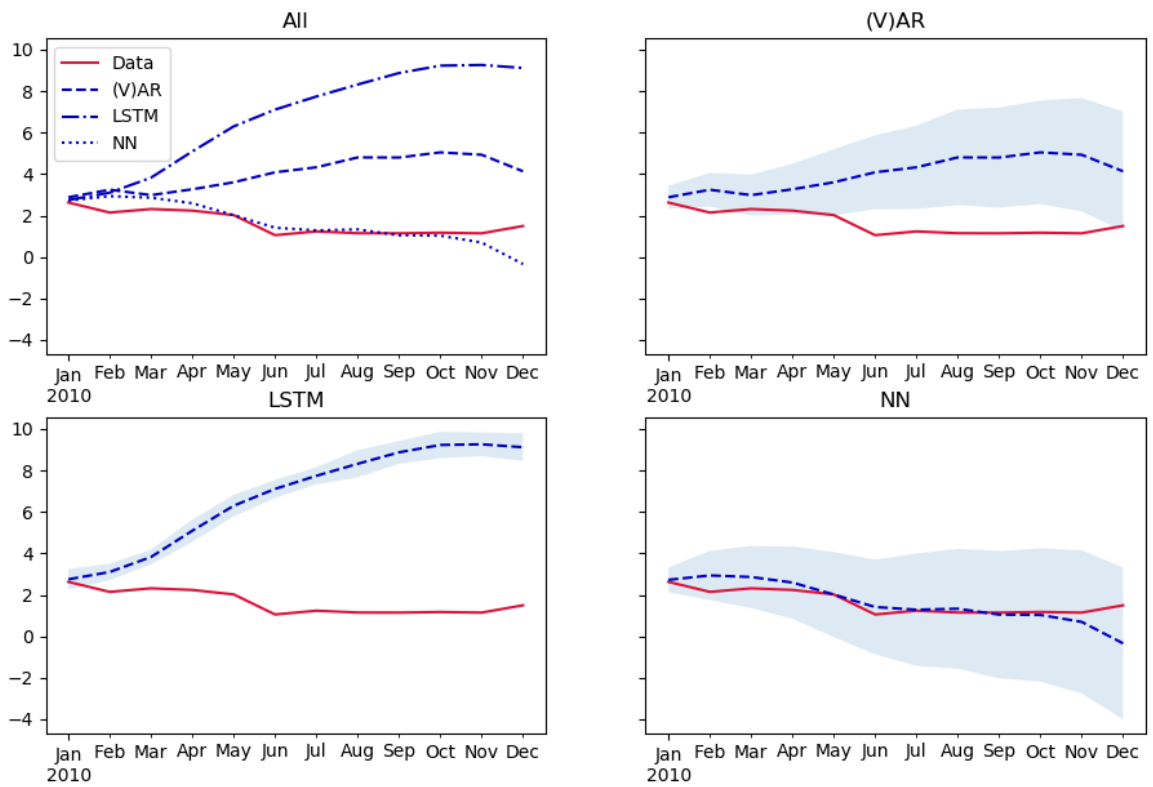
**Figure 2.7:** Real-Time Forecast 2007



**Figure 2.8:** Real-Time Forecast 2008



**Figure 2.9:** Real-Time Forecast 2009



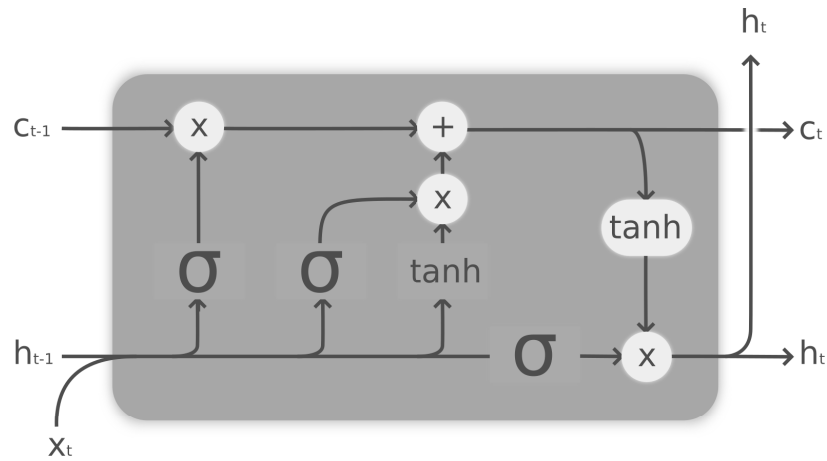
**Figure 2.10:** Real-Time Forecast 2010

to capturing the nonlinear features of the data, these recurrent neural networks are able to extract the relevant information about the trend and seasonal components of the data without a prior data preprocessing. We further investigate the real-time forecasting performance of the LSTM by computing 1-month-ahead predictions in a rolling-window setting. We find that its prediction is significantly superior to the NN and AR forecasts. However, the errors of indirect LSTM forecasts at longer horizons exhibit high variation. LSTMs are extremely accurate in some years and produce "wild forecasts" in others. Implausible forecasts can presumably be eliminated by the researcher through forecast trimming.

In the light of our results further analysis of the LSTM's performance on predicting other macroeconomic time series seems to be a promising avenue for future research.

## 2.8 Appendix

### LSTM Representation



**Figure 2.11:** LSTM Cell Representation

$x_t$  - input at step  $t$

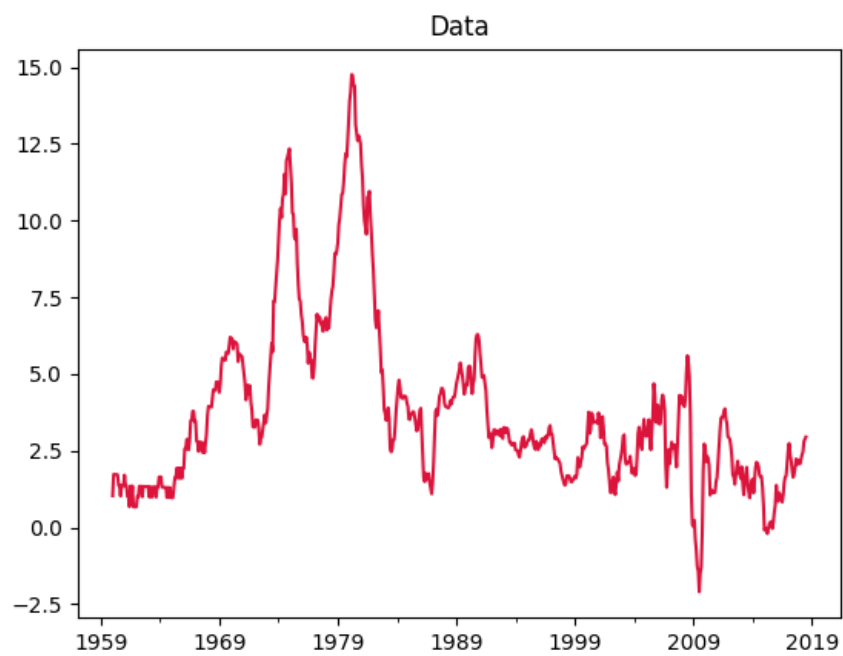
$c_t$  - memory (state) at step  $t$

$h_t$  - output at step  $t$

$\sigma$  - sigmoid function

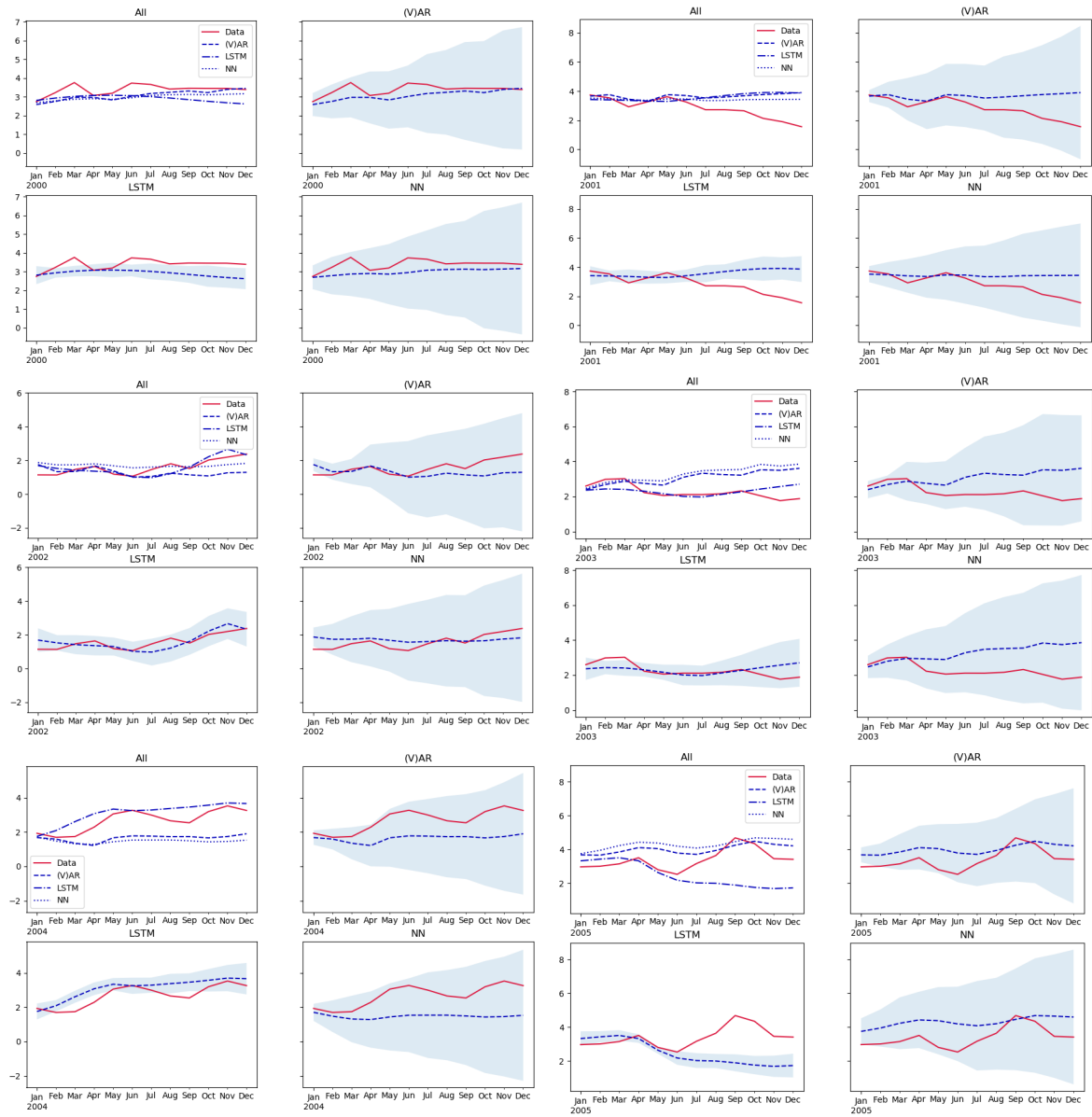
tanh - hyperbolic tangent function

## CPI Inflation Data



**Figure 2.12:** Raw CPI Inflation Data

## 12-Months-Ahead Forecasts



**Figure 2.13:** Real Time Forecasts, Forecast Step  $h=1:12$ , 2000-2005

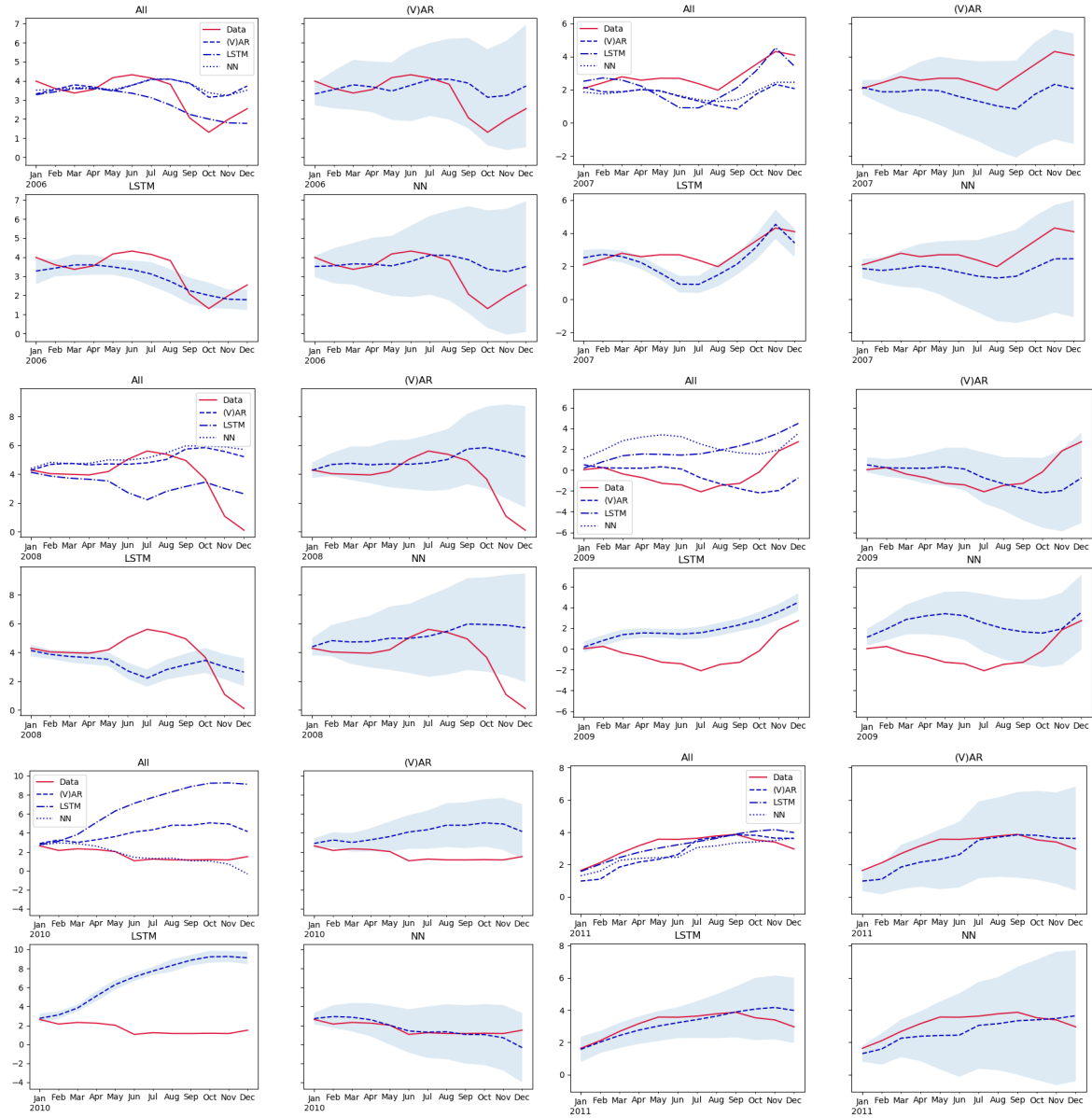
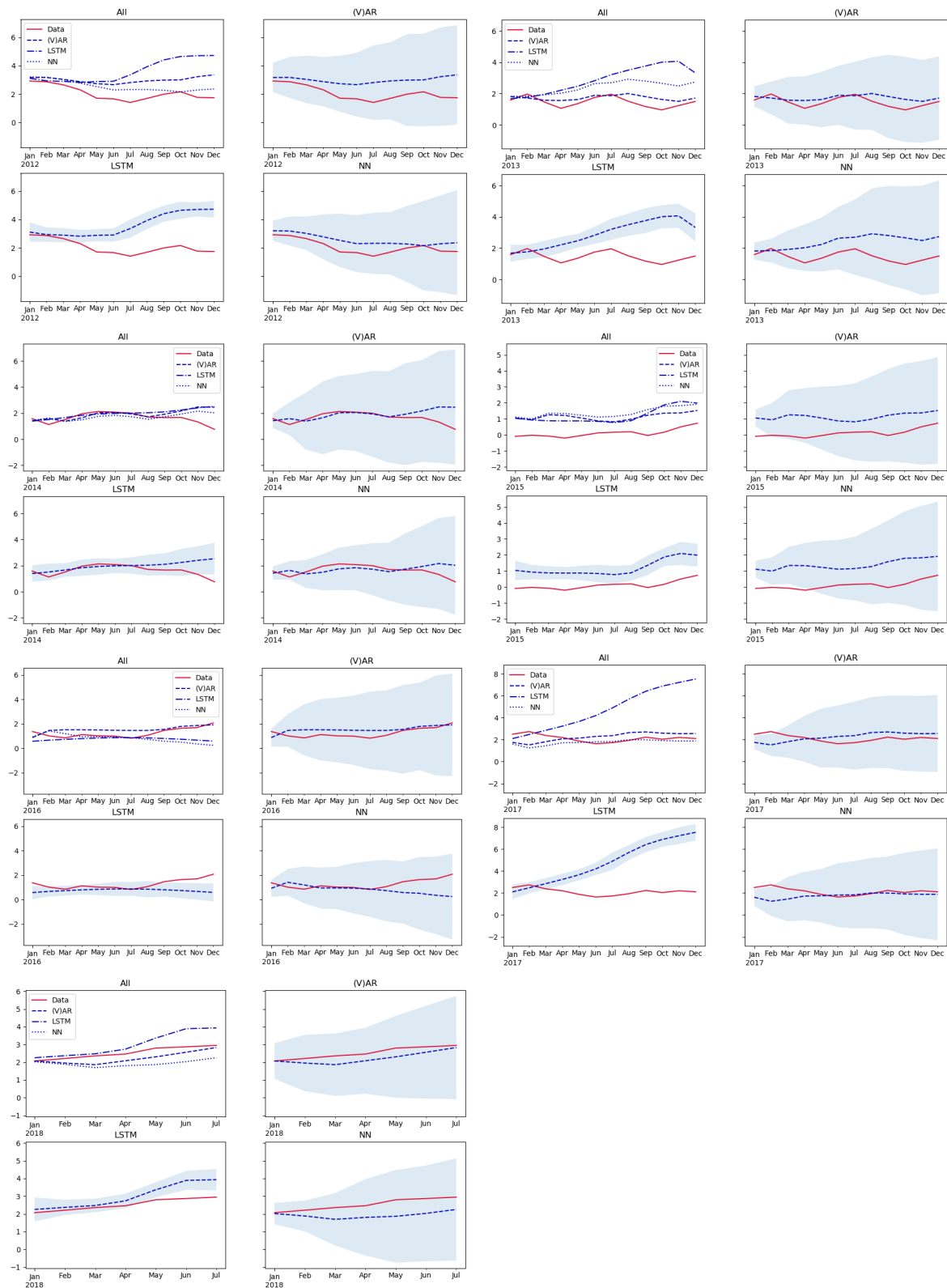


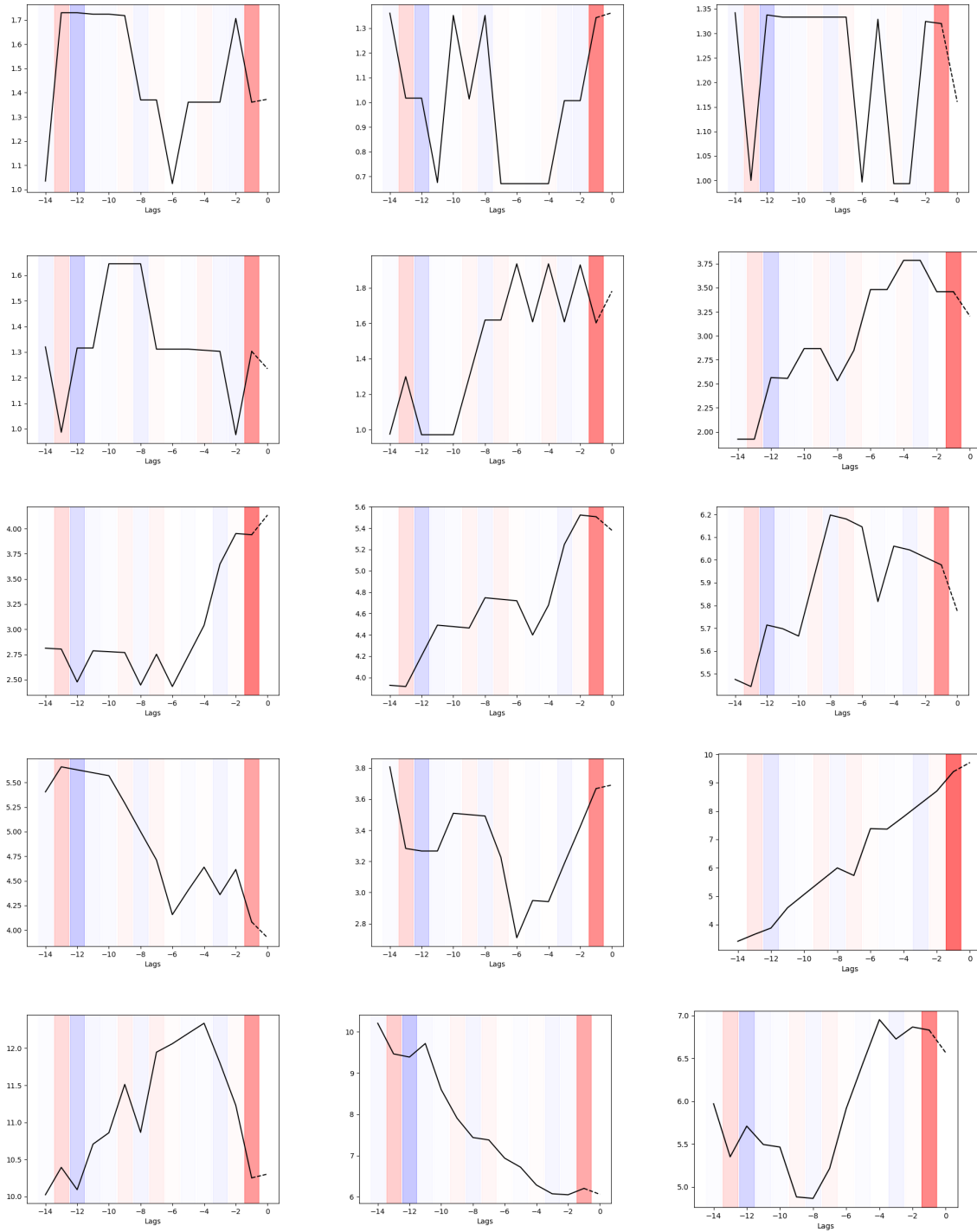
Figure 2.14: Real Time Forecasts, Forecast Step  $h=1:12$ , 2006-2011



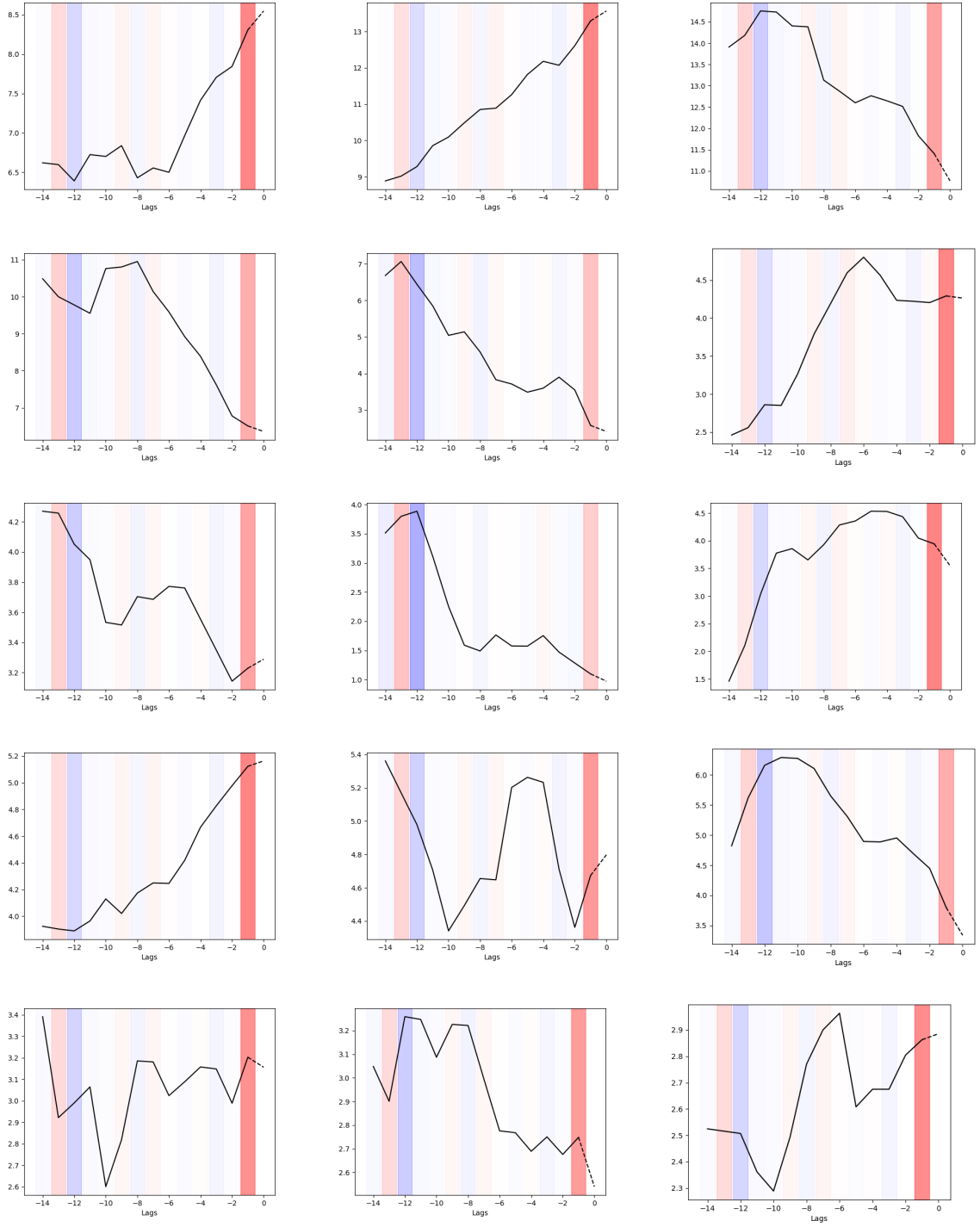


**Figure 2.15:** Real Time Forecasts, Forecast Step  $h=1:12$ , 2012-2018:07

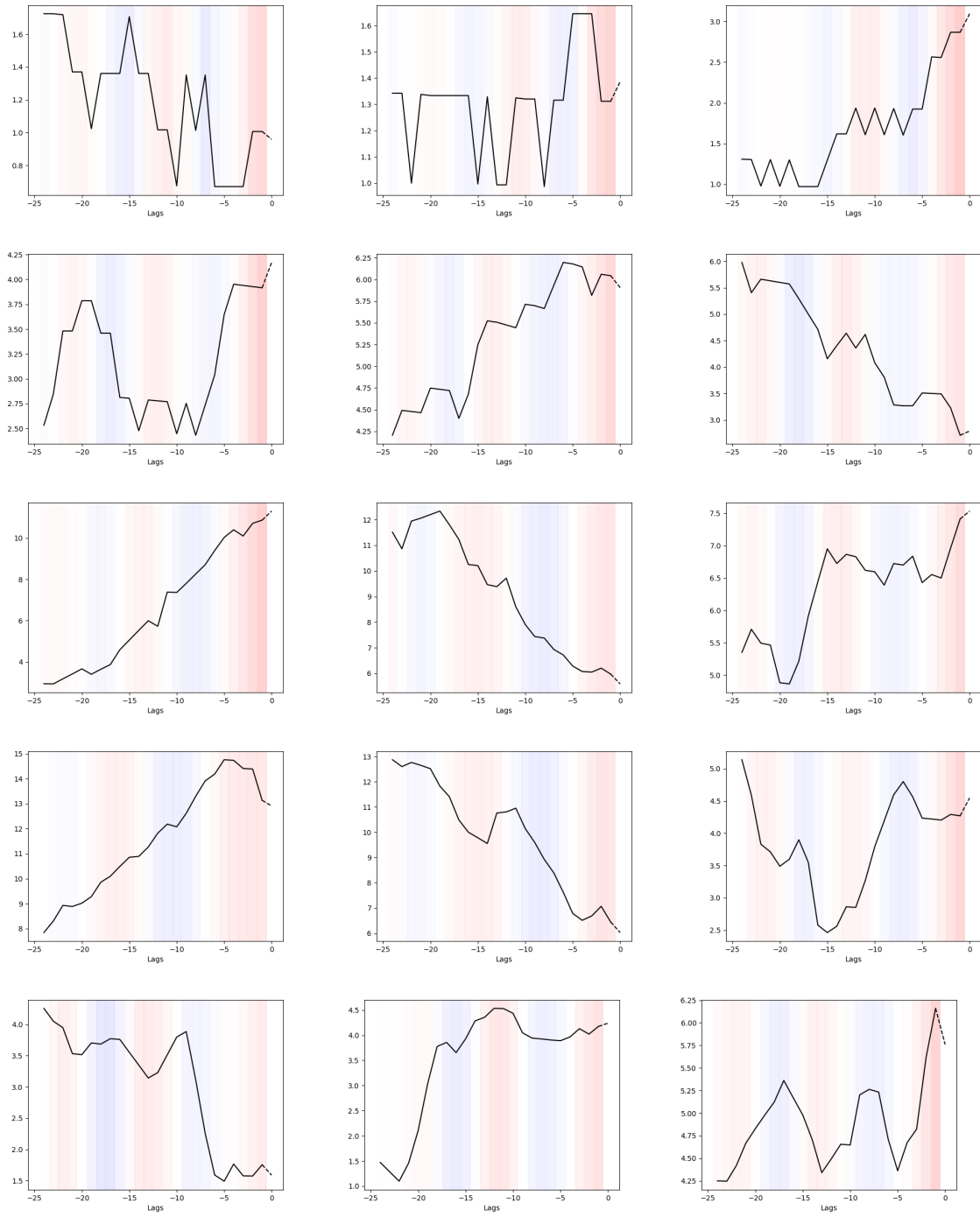
# LRP Analysis



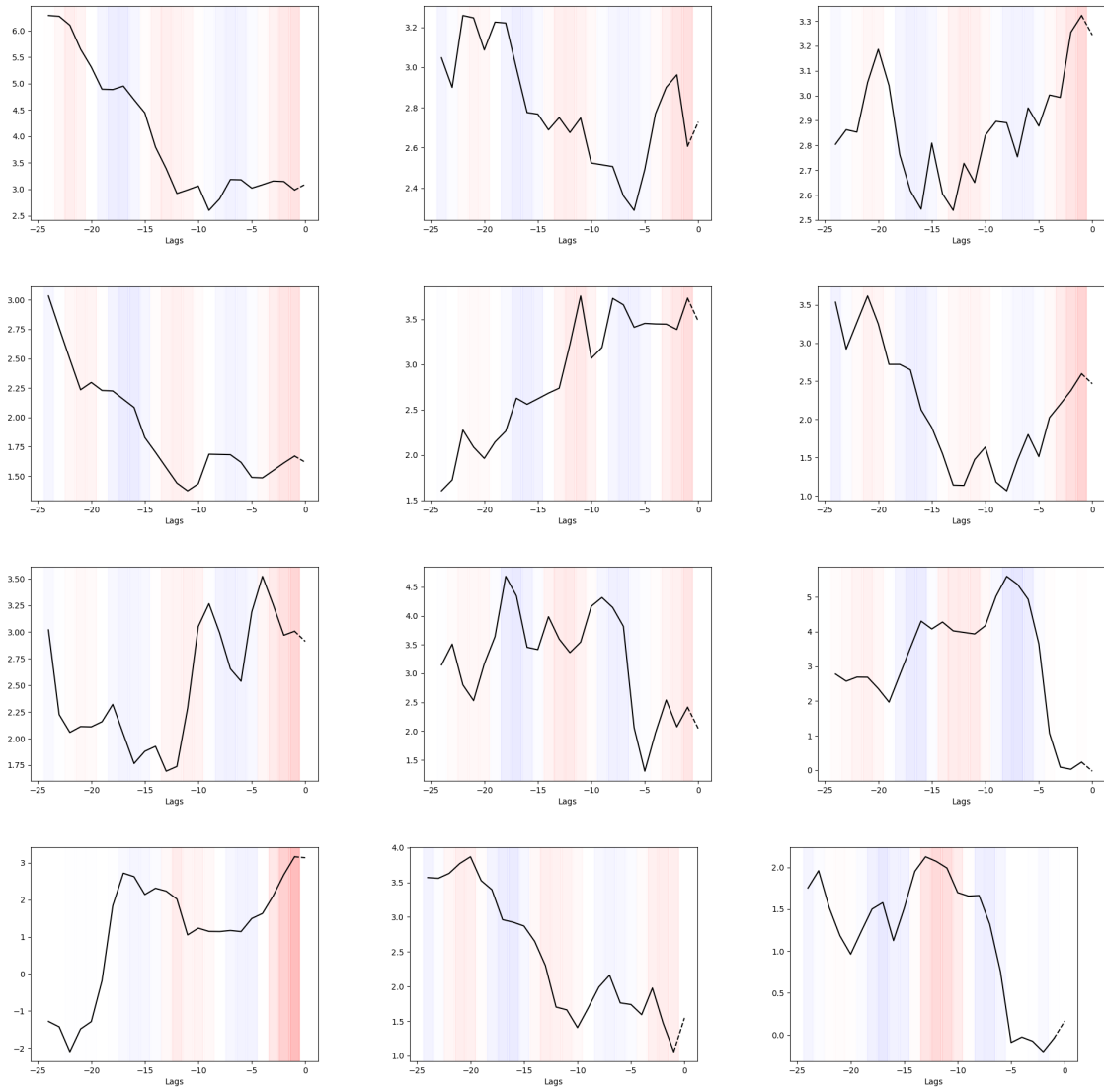
**Figure 2.16:** Layer-wise Relevance Propagation Analysis for Simple NN, Examples (part 1)



**Figure 2.17:** Layer-wise Relevance Propagation Analysis for Simple NN, Examples (part 2)



**Figure 2.18:** Layer-wise Relevance Propagation Analysis for LSTM, Examples (part 1)



**Figure 2.19:** Layer-wise Relevance Propagation Analysis for LSTM, Examples (part 2)











### 3 Asymmetries in Interest Rate Dynamics

#### Abstract

This paper analyzes if the reaction of the monetary policy to inflation and output growth is asymmetric around the targets for these variables. It asks if the central bank fights too high inflation as actively as too low inflation and if it is concerned about too low output growth as much as about too high output growth. The question is analyzed in a nonlinear New-Keynesian DSGE model. I introduce a flexible formulation of the Taylor rule which allows for a state- dependent asymmetry in its coefficients. I then let the data determine the degree of this asymmetry by estimating the model on US data for 1984-2007. I estimate the second-order approximation of the solution of the DSGE with the help of the particle filter and obtain the posterior draws by the Random Walk Metropolis-Hastings algorithm. I find significant asymmetries in the reaction function of the US central bank. The reaction to inflation is stronger when inflation is higher than the target and the reaction to output growth is stronger when the output growth rate is below the target. I further show that despite the identified nonlinearities the interest rate dynamics in the model fails to pass a second-order accuracy test that was suggested by Aruoba, Bocola, Schorfheide (2017).

**Keywords:** Interest rate rules, Nonlinear DSGE, Particle filter, Sequential importance sampling

**JEL classification:** C11, C32, E32, E43

#### 3.1 Introduction

Dynamic stochastic general equilibrium models (DSGE) became a workhorse of macroeconomic research and policy analysis in academia as well as in central banks. The behavior of monetary policy in these models is usually formulated as a linear Taylor (1993) rule. According to this rule the nominal interest rate reacts with some positive coefficient to movements of inflation and with some positive coefficient to the movements of output or the output growth rate. The reaction coefficients are independent of the sign and size of inflation and output deviations from their targets. While being a good first-order approximation of the actual policy decisions, linear Taylor rule might be too much of a simplification especially if one analyzes macroeconomic dynamics in a nonlinear model.

Several papers conclude that the actual behavior of the central banks is state dependent. Reaction of the central banks is not symmetric across positive and negative deviations of its target variables. A model that ignores these asymmetries would provide misleading

conclusions about the propagation of the macroeconomic shocks and would not be suitable for developing policy recommendations.

Aruoba et al. (2017) suggest a formal procedure that allows to evaluate asymmetries in a DSGE model. They show that a standard New-Keynesian DSGE with a linear Taylor rule is unable to generate any degree of the asymmetry in the interest rate dynamics. At the same time they document large asymmetries in the movements of federal funds rate in the data during 1984-2007. As the authors state the linear monetary policy rule might be the reason for the symmetric interest rate behavior in the model: *"Since the interest-rate feedback rule does not feature any asymmetries, the policy instrument does not display the asymmetry we identified in the data"*.

This paper moves beyond the standard formulation of the Taylor rule. It studies if the reaction of the monetary authority to inflation is asymmetric across positive and negative deviations of inflation from its target and whether the reaction to output growth rate is asymmetric across positive and negative deviations of the output growth rate from its target. It also investigates if a nonlinear formulation of the monetary policy can reconcile the interest rate dynamics in the model and in the data. I introduce a flexible formulation of the Taylor rule into a New-Keynesian DSGE model. The suggested rule is a sum of two independent linear-exponential (linex) functions - one is the function of inflation deviations from its target and the other is the function of the output growth rate deviations from its target. This formulation leaves the possibility that the reaction of the monetary policy is linear for positive deviations and exponential for negative deviations or the other way around. The exact shape and the sign of the asymmetry is determined by the coefficients in the linex function.

I evaluate empirical evidence for the asymmetry in the interest rate rule by bringing the model to the data. I run a Bayesian estimation of the model on US data for 1984-2007. I do not use the data after 2007 because I want to avoid the episode when the "zero-lower bound" on the interest rate was binding. Since the model is nonlinear I employ a sequential importance sampling (particle filter) for the likelihood approximation. The draws from the posterior distributions of the parameters are obtained by the Random Walk Metropolis-Hastings. Nonlinear estimation is essential to my analysis. In a log-linearized model all policy functions are symmetric across the states by construction.

On the next step, I proceed to access the interest rate dynamics in the model. I run posterior-predictive checks that were developed in Aruoba et al. (2017) for second-order DSGE models. The authors suggest to estimate a univariate quadratic autoregressive model

(QAR) on the actual data and on the data that was simulated from a DSGE model. Differences between the QAR coefficients would measure the discrepancy between the data and the model.

My findings can be summarized as two main results. First, my estimation confirms that significant asymmetries are present in the interest rate dynamics. The US central bank increases nominal interest rate more aggressively (exponentially) when inflation is above the target. When inflation is below the target the reaction of the interest rate is linear. The overall degree of the asymmetry in the reaction to inflation is moderate. For the output growth rate asymmetries are more pronounced. The interest rate is increased linearly when the output growth is above the target and is reduced exponentially when the economic growth is too weak. Posterior odds ratio tests indicate that the model with the asymmetric monetary rule is closer to the data.

Second, I show that the current model with the asymmetric interest rate is unable to satisfy the QAR posterior predictive checks. The movements of the interest rate in the model are symmetric. In other words, interest rate increases are as likely as interest rate declines. One possible explanation for this result might be that the model generates inaccurate dynamics of the central bank targets. I test this hypothesis by running the following counter-factual exercise. I insert the actual data for the GDP growth rate and inflation into the estimated asymmetric monetary policy rule and compute counter-factual interest rate data. I then use this data to estimate the coefficients of the QAR model. This experiment lends no support to the hypothesis that the model would be able to reproduce asymmetric interest rate dynamics if it had perfectly replicated the dynamics of the central bank targets. I further show that nonlinearities in the interest rate rule play little role in shaping the response of the model economy to macroeconomic shocks. Generalized impulse response functions show no significant dependency on the state of the economy or the size and sign of the shock. The contribution of the paper is documenting that asymmetries in the reaction of the monetary policy alone cannot account for the observed asymmetric interest rate dynamics.

This paper is not alone in analyzing asymmetric preferences of the central banks. A variety of the empirical studies conclude that interest rates behave differently in recessions compared to booms and in high inflation regimes compared to low inflation regimes. For example, Kesriyeli et al. (2004) conducted a reduced form analysis for the US, Germany and the UK and concluded that different coefficients are applicable for the monetary policy rules when interest rates are increasing versus when they are decreasing. Qin and Enders

(2008) fitted a logistic and an exponential function to the US Federal Funds rate. They show that the data generally prefers a nonlinear rule over the linear one. Martin and Milas (2004) considered the reaction of the Bank of England and found that the policy responds more to positive inflation deviations from the target. Cukierman and Muscatelli (2008) estimate a Taylor rule for the US and the UK by a smooth-transition regression. They found a higher reaction to inflation of the US central bank during the Vietnam War, the Burns and Miller and the Greenspan periods. Dolado et al. (2004) find sign and size asymmetries in the US Taylor rule after 1983. Markov and de Porres (2012) demonstrated with a non-parametric estimation that in 8 OECD countries central banks use a significantly nonlinear Taylor rule. The largest asymmetry appears along the size of the current inflation level.

The literature suggests that the asymmetries in monetary policy rules can be estimated in different ways (for example, with a threshold regression, smooth transition regression, regime switching model, with the help of time varying coefficients or time-varying targets of the central bank). Examples include Bunzel and Enders (2010), Owyang and Ramey (2004), Cogley and Sargent (2005), Boivin (2006), Sims (1999), (2001), Sims and Zha (2006), Assenmacher-Wesche (2005), Alcidi et al. (2005 and 2009), Gerlach and Lewis (2014) and Bruegermann and Riedl (2011) among others.

The above-mentioned literature focuses on either a reduced form estimation or SVARs. In contrast, this paper estimates the interest rate asymmetries in a structural macroeconomic model. Estimation of a fully specified DSGE model reduces the bias of the estimates since the model generates cross-equation restrictions from the structure of the economy and thus accounts for the endogeneity problem (Finocchiaro and von Heideken, 2013). Moreover, as discussed in Lubik and Schorfheide (2007) parameters are better identified because the structural model forces the volatility of co-movements of the variables to be consistent with the data. Another aspect of this study which is novel to the literature is that, to the best of my knowledge, this is the first paper that studies whether an asymmetric Taylor rule can help to reconcile the DSGE model with the evidence on asymmetries in the interest rate data.

It turns out that the assessment of the DSGE's ability to reproduce the nonlinear features of the data is not a trivial task. While an extensive literature discusses the comparison of linear DSGEs and VARs (for example, Schorfheide, 2000; Del Negro and Schorfheide, 2004; Smets and Wouters, 2003 and 2007), the methods for nonlinear models are rather scarce. Some examples are Aruoba et al. (2017), Barnichon and Matthes (2016), Ruge-Murcia (2016). I employ a novel approach of Aruoba et al. (2017) that was developed precisely for the

nonlinear DSGEs. This procedure is analogous to the posterior predictive checks that are used for the linear models.

The analyses in this paper also shed some light on whether an asymmetric monetary rule affects the propagation of the shocks in the model. The literature on non-linear VARs finds that impulse responses are asymmetric across the states of the business cycle (see for example, Weise, 1999; Barnichon and Matthes, 2016). At the same time, the state dependency of impulse responses in a DSGE model approximated up to the second order is rather insignificant, for example Amisano, Tristani (2010) or Andreasen et al. (2017). This result is preserved in the current model. I plot the generalized impulse response functions (GIRF) after a technology shock and after a monetary policy shock for 1) different states of the economy and 2) for different signs and magnitudes of the shocks. 95% credible sets of the impulse responses almost coincide. Consequently, one cannot reject the hypothesis that these GIRFs are statistically indistinguishable.

Although this paper does not comment upon justifications for the asymmetry in monetary policy rules it is complementary to the literature that works on this topic. For example, one of the reasons suggested in the literature is a nonlinearity of the aggregate supply (due to the presence of adjustment costs). The nonlinear Phillips curve implies that the central bank faces a different inflation-output trade-off in recessions than in booms. Dolado et al (2004, 2005), Latxón et al. (1995, 1999), Alvarez-Lois (2000) among others provide some empirical evidence for a convex Phillips curve in several European countries and the US. Another explanation might be that the central bank has asymmetric preferences over recessions versus booms and high inflation versus low inflation regimes (see Orphanides and Wieland, 2000; Ruge-Murcia, 2002; Dolado et al., 2004; Surico, 2002 and Cukierman and Muscatelli, 2002). For example, Cukierman and Muscatelli (2002) find empirical evidence for a precautionary demand of policymakers for expansions and for low inflation in the US, Japan and the UK which is in accordance with my results for the US. Finally, the way the expectations are formed might be sensitive to the current economic conditions which leads to a nonlinear optimal policy rule as discussed in Meyer et al. (2001).

The rest of the paper is organized as follows. Section II presents empirical evidence on the asymmetric interest rate dynamics. Section III sketches the model and introduces an asymmetric interest rate rule. Section IV provides the details on the model estimation and the generalized impulse responses. Section V presents posterior predictive checks based on QAR. Section VI concludes.

## 3.2 Evidence on Asymmetry in Interest Rate Dynamics

This section documents empirical evidence on the asymmetry in interest rate movements. First, it presents empirical observations on asymmetry in the changes of the Federal Funds rate - negative changes are less frequent but are larger in absolute terms than the positive changes. Second it argues that the standard linear Taylor-type reaction function cannot capture the asymmetric behavior of the interest rate. Finally, it discusses the findings of Aruoba et al. (2017) that a DSGE model with a standard Taylor rule is unable to generate plausible degree of asymmetries in the interest rate movements.

Throughout the paper I use US data from the FRED database of the Federal reserve Bank of St. Louis. I use data on the nominal interest rate, real GDP and the GDP growth rate, nominal wage growth and inflation. Real GDP growth rate is obtained as log differences of real GDP. To compute per capita real GDP growth the smoothed population growth rates were used as in Aruoba et al. (2017). Inflation was computed as the log differences of the GDP deflator. Nominal wage growth is the log difference of compensation per hour in the non farm sector. Quarterly data for the nominal interest rate is computed from the monthly averages of the Fed funds rate. The DSGE estimation is conducted on quarterly series from 1984 to 2007. This time span is motivated by two reasons. First, as I show in the next chapter for quarterly data the largest asymmetry in the interest rate dynamics is present in the after Volker-era that is in the post-1983 samples. Second, I do not consider observations after 2007 since I focus on normal times and not on the zero-lower-bound episode. The asymmetry of the interest rate that arises from the zero-lower-bound constraint is not the focus of this study.

### 3.2.1 Interest Rate Dynamics

Figure 3.1 presents monthly data for the US effective Federal funds rate for 1960-2016. Shaded areas indicate NBER recessions. Visual inspection of the data suggests that after the end of the Volcker era, that is for the post-1983 period, declines of the interest rate were on average of a larger magnitude than increases but less frequent. A decline of the interest rate is normally followed by a sequence of small positive or near-zero interest rate increments.

Table 3.1 shows summary statistics for the monthly changes in interests rate for the sample from 1960 to 2007, as well as for pre-1983 and post-1983 sub-samples.

Several observations emerge. In all periods the number of interest rate declines is smaller than the number of periods when the interest rate was increased or unchanged. At the same



**Figure 3.1:** Monthly Effective Federal Funds Rate, 1960:01-2018:10. Source: FRED

**Table 3.1:** Summary Statistics of Interest Rate Changes, %

	Mean	Negative		Non-negative	
		amount	largest	amount	largest
1960-2007	0.00	250	-6.63	325	3.06
1960-1983	0.02	123	-6.63	164	3.06
1984-2007	-0.02	127	-1.31	161	0.87

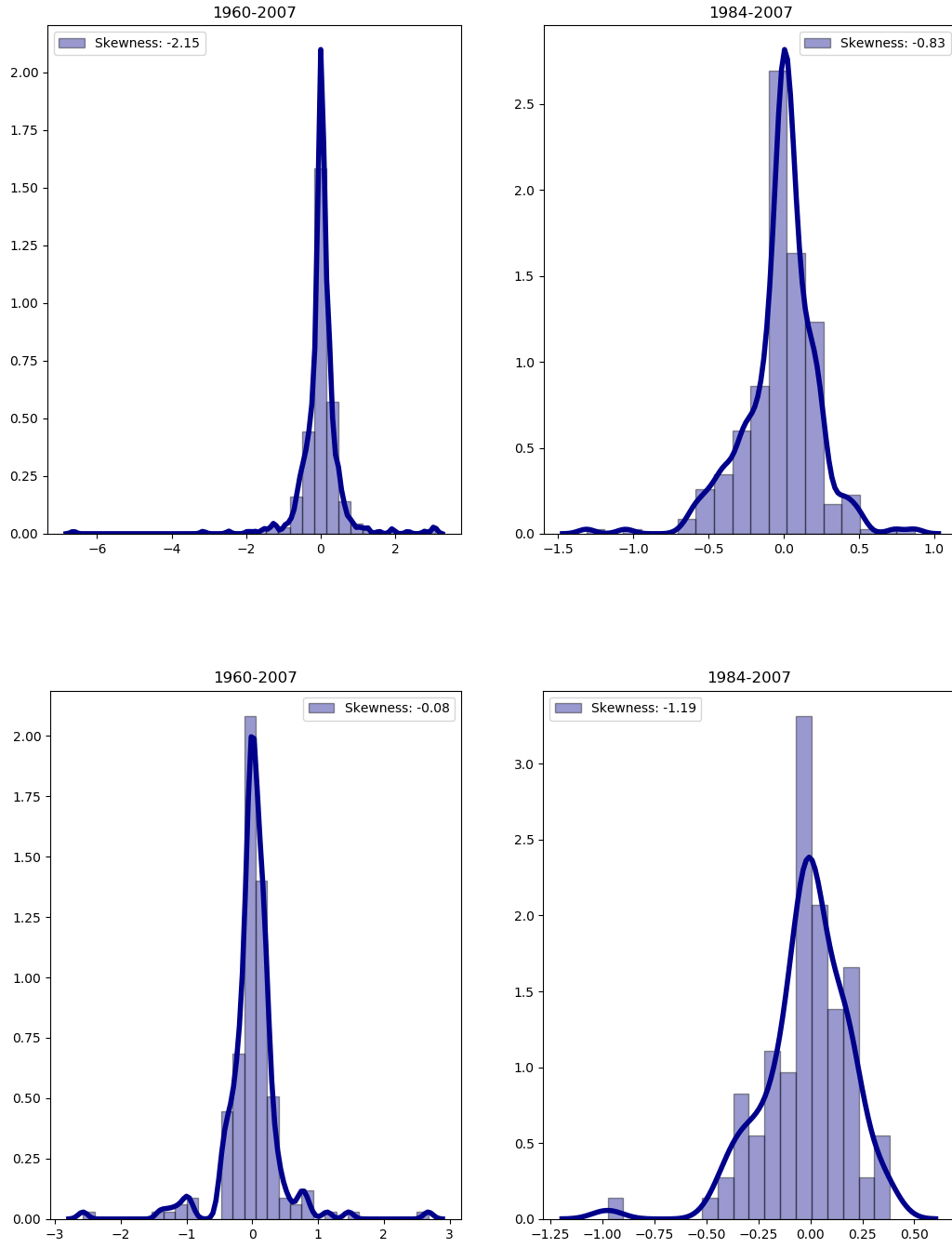
time, the largest decrease in absolute terms exceeds the largest increase and the mean change is nearly zero. It means that fast declines in the interest rate are followed by several moderate increases. In other words, the Federal Reserve is ready to lower the interest rate by a large amount but seems to be more reluctant to drastically increase it.

Figure 3.2 confirms this observation by presenting density plots of the changes in the interest rate during the whole sample as well as during the 1984-2007 period. The distribution of the interest rate movements is left skewed. This result is present in both monthly (top row) and quarterly data (bottom row).<sup>13</sup> The skewness coefficient is negative for all considered samples.

At least two reasons might underline these observations. On the one hand, targets of the central bank might display asymmetric behavior. If GDP drops sharply in recessions and recovers slowly and gradually afterwards then this asymmetry would translate into the interest rate dynamics. On the other hand, the central bank might have asymmetric preferences and, therefore, an asymmetric interest rate rule. If the Federal Reserve is recession-averse then it responds to output declines by immediate sharp interest rate decreases and to output rises

<sup>13</sup>Quarterly data were obtained by taking quarterly averages of the monthly data.





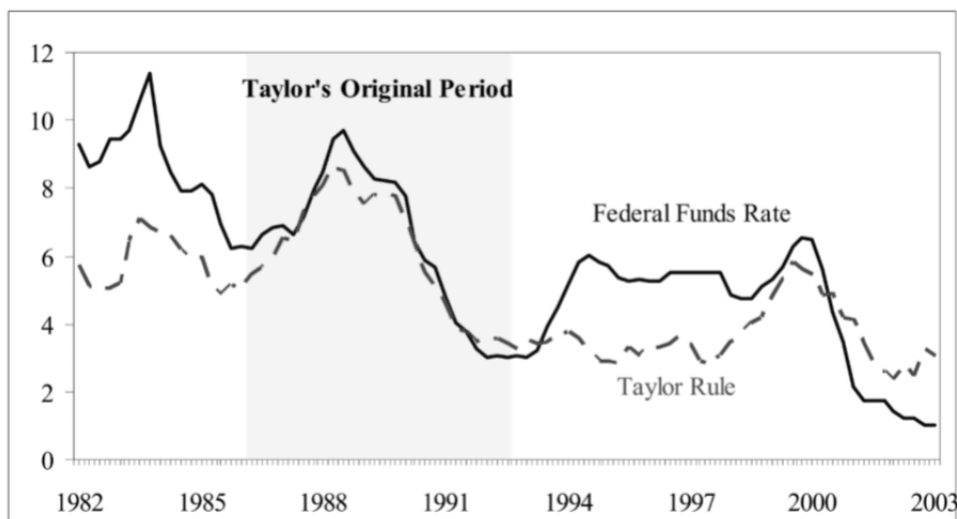
**Figure 3.2:** Histogram and Kernal Density Estimation for Changes in the Interest Rate.

*Note.* Monthly (top row) and quarterly (bottom row) frequency is presented.

by a series of graduate interest rate increments. The next part of this section shows that asymmetry in the central bank targets cannot fully explain the observed skewed distribution of the interest rate adjustments.

### 3.2.2 Linear Monetary Policy Rule

The vast majority of DSGE models include a linear monetary policy rule from the influential work of John Taylor (1993). While being a good approximation of the monetary policy this rule might be subject to some criticism. For example, Rabanal (2004) points out that this rule achieves a satisfactory goodness of fit only for 1986-1993, the years that were used to derive the rule in the original paper (Figure 3.3).



**Figure 3.3:** Taylor Rule versus Real Data, US, 1982-2003. Source: Rabanal (2014)

As a first check for nonlinearity I estimate two versions of a linear Taylor rule - one with inflation and output gap and one with inflation and output growth rate.<sup>14</sup> Appendices (3.7) and (3.7) display the plots of residuals from both estimated Taylor rules. In the second part of the sample, 1984-2012, Taylor residuals are negatively correlated with the GDP growth rate with the correlation coefficient being -0.34 (and -0.0031 for the output gap). This means that the linear specification of the interest rate rule fails to capture a part of the variation in the Federal funds rate that can be explained by the movements of the GDP growth rate.

More importantly, the linear Taylor rule cannot reproduce the asymmetries in the interest rate dynamics that were discussed in the previous subsection. Figure 3.4 presents the density plots for artificial interest rate data computed with two linear Taylor rule: the linear Taylor rule with standard coefficients and with or without interest rate smoothing, that is either  $R_t = 1.5\pi_t + 0.5y_t$  or  $R_t = 0.8R_{t-1} + 0.2 \cdot (1.5\pi_t + 0.5y_t)$ . The target inflation rate and the target GDP growth rate do not matter for the changes in the interest rate and are, therefore,

<sup>14</sup>The output gap was computed by HP-Filter with the smoothing parameter 1600.

omitted. I then take the first differences of the artificial interest rate data and produce density plots.

The resulting densities are highly symmetric and the skewness coefficients are small and positive for all considered cases. This result also holds for the Taylor rule that includes the output gap which is defined as actual GDP minus potential GDP. It indicates that asymmetry in the interest rate adjustments cannot be fully explained by asymmetries in the inflation and/or output dynamics. Asymmetric reaction of the monetary policy is at least partially responsible for the observed regularity.

### 3.2.3 QAR Analysis

Aruoba, Bocola, Schorfheide (2017) document similar observations after estimating a quadratic autoregressive time series model (QAR) on US data. The main idea of the method is to represent the solution of a DSGE model as  $y_t = f(y_{t-1}, \omega u_t)$ . This solution can be expressed as  $y_t^* = y_t^{(0)} + \omega y_t^{(1)} + \omega^2 y_t^{(2)}$  where  $y_t^{(1)}$  collects first order terms and  $y_t^{(2)}$  second order terms (Lombardo, 2010). The second order approximation of this solution around the point  $(y_t = y^*, \omega = 0)$  can be rewritten as a nonlinear state space model - quadratic autoregressive model (QAR):

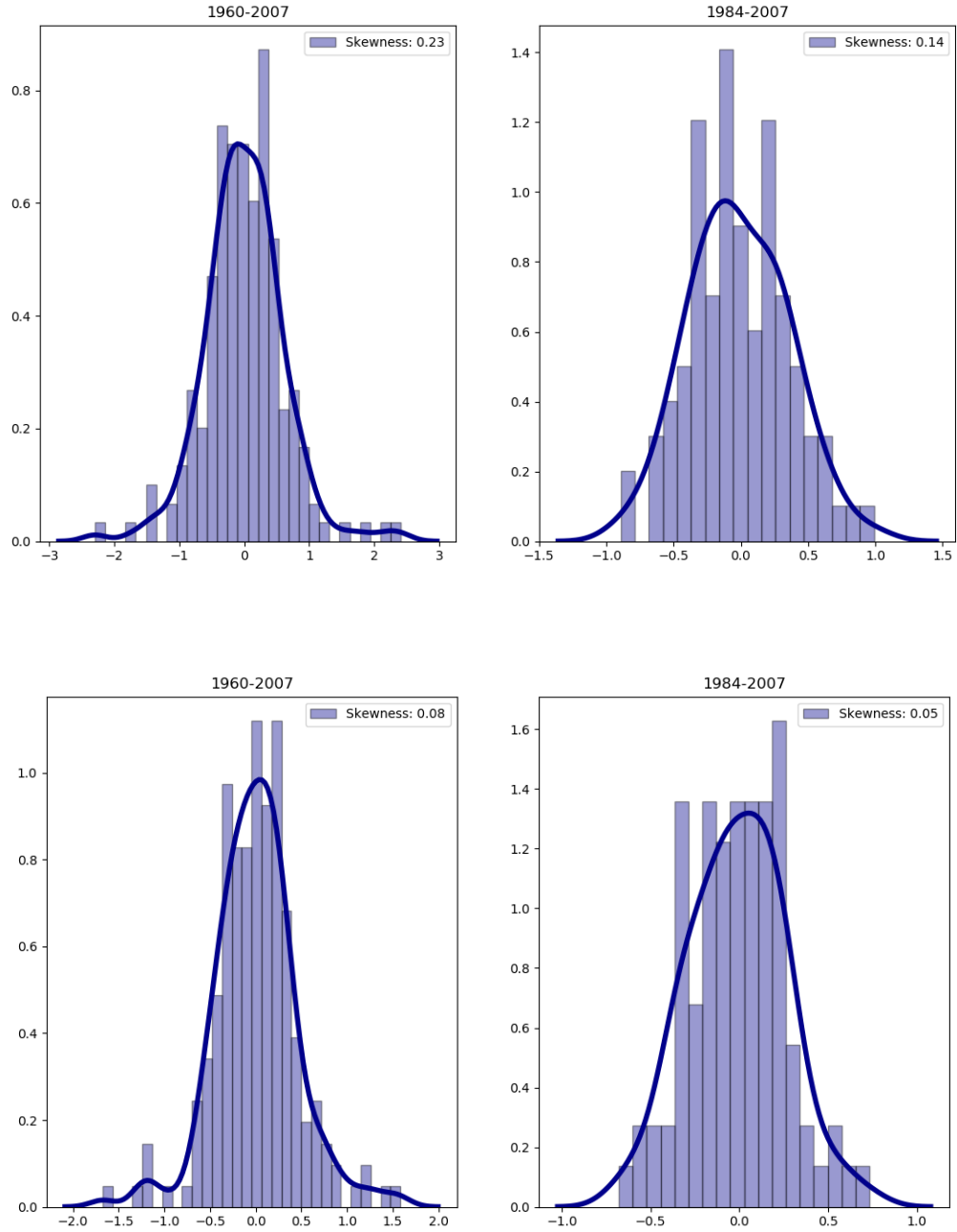
$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + \phi_2 s_{t-1}^2 + (1 + \gamma s_{t-1})\sigma u_t \quad (3.1)$$

$$s_t = \phi_1 s_{t-1} + \sigma u_t, \quad u_t \text{ iid} \sim N(0, 1) \quad (3.2)$$

where  $y_t$  is a variable of interest,  $\phi_0$  is the unconditional mean,  $\phi_1$  is an autoregressive coefficient, the coefficient  $\gamma$  estimates conditional heteroscedasticity and the coefficient  $\phi_2$  captures asymmetries of the time series.  $s_t$  is an unobserved state which itself follows an AR(1) process. The moments and the IRF functions of this model are state dependent.

The QAR model allows to jointly estimate asymmetry and heteroscedasticity. Moreover, it closely resembles a second-order solution of a DSGE model and can, therefore, be used to evaluate its second-order accuracy. The authors suggest QAR-based posterior predictive checks that use the same procedure as VAR-based posterior predictive checks for linear models and contain two steps: 1) fit a QAR time-series model on the actual data and the data that were simulated from a DSGE, 2) compare the QAR coefficients: similar coefficients indicate a good performance of the DSGE model.

For the purpose of this study, two main conclusion of Aruoba et al. (2017) are important. First, the authors show that in the post-1984 samples, the  $\phi_2$  coefficient for the Federal funds rate is significantly negative. For the 1984-2007 subsample,  $\phi_2 = -0.16$ . This means that



**Figure 3.4:** Histogram and Kernel Density Estimation for Changes in the Interest Rate under a linear Taylor Rule.

*Note:* The level of quarterly interest rate was computed from the Taylor rule without interest rate smoothing:  $R_t = 1.5\pi_t + 0.5y_t$  (top row) and with smoothing:  $R_t = 0.8R_{t-1} + 0.2 \cdot (1.5\pi_t + 0.5y_t)$  (bottom row). The changes were computed as first differences.

$\phi_2 s_{t-1}^2 \leq 0$  across all states, which implies that interest rate falls faster than it rises. Second, they found a large discrepancy between the asymmetry that is present in the actual interest rate data and the asymmetry in the interest rate dynamics in a standard New Keynesian DSGE model. The authors conclude that a linear Taylor rule might be responsible for this result. In the next section I present the model that was used in the Aruoba et al. (2017) and investigate if a nonlinear Taylor rule specification can help to reconcile the model with the data.

### 3.3 The Model

The DSGE model of the economy is identical to Aruoba et al. (2017) except for formulation of the interest rate rule. This choice is motivated by three main reasons. First the model is a version of Smets and Wouters (2007) New Keynesian DSGE which became a workhorse in macroeconomic analysis and economic policy design. The model features a balanced growth path (BGP) which is introduced in the form of a labor augmenting technology growth. This feature is beneficial for the estimation since it allows to use non-detrended data. Second the model is augmented by asymmetric wage and price adjustment costs as in Kim, Ruge-Murcia (2009). The degree of the asymmetry in each case is driven by a pair of the parameters. These parameters can easily be changed in order to address the effect of the wage and price nonlinearities on the interest rate dynamics. Finally, the results of the model estimation can be directly compared with the findings in Aruoba et al. (2017).

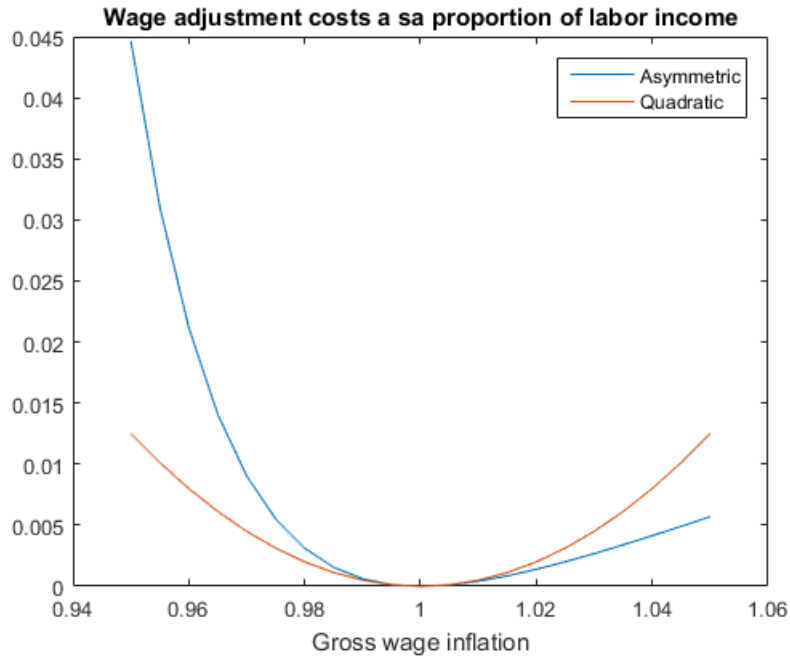
The economy is populated by a continuum of heterogeneous workers. Every worker  $k$  optimizes her utility (3.3) from the specific consumption level  $C(k)$  relative to the current technology level minus her disutility from work. All workers live in one representative household and are perfectly insured against each other by "sharing the same table".

$$\mathbb{E}_t \left[ \sum_{s=0}^{\infty} \beta^s \left( \frac{(C_{t+s}(k)/A_{t+s})^{1-\tau} - 1}{1-\tau} - \chi_H \frac{H_{t+s}^{1+1/\nu}(k)}{1+1/\nu} \right) \right] \quad (3.3)$$

Every household member sets a wage for her specific labor of type  $k$  and bears wage adjustment costs (3.4). These costs are expressed as a share of nominal wage income.  $\gamma\pi$  is a growth rate of the nominal wage at the BGP,  $\gamma$  is a technology growth rate and  $\pi$  is a rate of inflation on the BGP.

$$\Phi_w \left( \frac{W_t(k)}{W_{t-1}(k)} \right) = \varphi_w \left( \frac{\exp(-\psi_w(\frac{W_t(k)}{W_{t-1}(k)} - \gamma\pi)) + \psi_w(\frac{W_t(k)}{W_{t-1}(k)} - \gamma\pi) - 1}{\psi_w^2} \right) \quad (3.4)$$

Wage setting costs are asymmetric which is modeled by a linear-exponential function which was introduced by Varian (1974). Under this functional form the wage adjustment costs depend not only on the magnitude of the wage change but also on the sign of adjustment. Figure 3.5 gives a graphical illustration. The degree of the asymmetry is controlled by the parameters  $\psi_w$  and  $\phi_w$ . The cost function is strictly convex for any  $\phi_w > 0$ . It nests a symmetric quadratic cost function as a special case when  $\psi_w$  approaches zero. For  $\psi_w > 0$  wage increases are associated with linear costs while wage decreases induce exponential costs. Hence, wages are more downward rigid. At the same time, wage decreases are not precluded and do happen rarely as observed in microdata (Kim, Ruge-Murcia, 2009).



**Figure 3.5:** Linex Adjustment Costs (Kim and Ruge-Murcia, 2009)

Household members buy government bonds  $B_t(k)$  which pay nominal interest  $R_t$ , receive profits from the firms  $D_t$  and pay lump-sum taxes  $T_t$ . The budget constraint for  $t \geq 0$  reads as

$$P_t C_t(k) + B_t(k) + T_t = W_t(k) H_t(k) \left( 1 - \Phi_w \frac{W_t(k)}{W_{t-1}(k)} \right) + R_{t-1} B_{t-1}(k) + P_t D_t(k) + P_t SC(k)_t$$

A labor aggregator aggregates all labor types via a CES aggregator with the elasticity of substitution between different labor types  $\lambda_w$ . Demand for each labor type depends negatively on the wage according to (3.5).

$$H_t = \left( \int_0^1 H_t(k)^{1-\lambda_w} dk \right)^{\frac{1}{1-\lambda_w}} \quad \text{where } H_t(k) = \left( \frac{W_t(k)}{W_t} \right)^{\frac{-1}{\lambda_w}} H_t \quad (3.5)$$

The nominal wage paid to workers is determined as (3.6).

$$W_t = \left( \int_0^1 W_t(k)^{\frac{\lambda_w-1}{\lambda_w}} dk \right)^{\frac{\lambda_w}{\lambda_w-1}} \quad (3.6)$$

Intermediate goods producers who are indexed by  $j$  operate under a monopolistic competition. They hire labor services  $H(j)$  and produce differentiated goods according to (3.7). Because of the monopolistic power, firms can set prices for their goods. Price adjustments are subject to costs defined by (3.8) as in Rotemberg (1982). The costs are expressed as a share of sales revenue and can be asymmetric in the sing of the adjustment. The degree of the asymmetry is controlled by the parameters  $\psi_p$  and  $\phi_p$ .

$$Y_t(j) = A_t H_t(j) \quad (3.7)$$

$$\Phi_p \left( \frac{P_t(j)}{P_{t-1}(j)} \right) = \varphi_p \left( \frac{\exp(-\psi_p(\frac{P_t(j)}{P_{t-1}(j)} - \pi)) + \psi_p(\frac{P_t(j)}{P_{t-1}(j)} - \pi) - 1}{\psi_p^2} \right) \quad (3.8)$$

$A_t$  is a level of productivity which follows an exogenous AR process integrated of order one:  $\ln A_t = \ln \gamma + \ln A_{t-1} + \ln z_t$  where  $\ln z_t = \rho_z \ln z_{t-1} + \epsilon_{z,t}$ .

All goods produced on the intermediate market are aggregated into a consumption basket by a final goods producer according to (3.9). As a result of the profit maximizing behavior the demand for an intermediate good  $j$  is negatively related to its price. The aggregated price level is given by (3.11).

$$Y_t = \left( \int_0^1 Y_t(j)^{1-\lambda_{p,t}} dj \right)^{\frac{1}{1-\lambda_{p,t}}} \quad (3.9)$$

$$Y_t(j) = \left( \frac{P_t(j)}{P_t} \right)^{\frac{-1}{\lambda_{p,t}}} Y_t \quad (3.10)$$

$$P_t = \left( \int_0^1 P_t(j)^{\frac{\lambda_{p,t}-1}{\lambda_{p,t}}} dj \right)^{\frac{\lambda_{p,t}}{\lambda_{p,t}-1}} \quad (3.11)$$

The elasticity of demand  $\lambda_{p,t} \leq 1$  is assumed to follow an AR(1) process  $\lambda_{p,t} = (1 - \rho_p) \ln \lambda_p + \rho_p \ln \lambda_{p,t-1} + \epsilon_{p,t}$  where  $\epsilon_{p,t}$  is white noise.

Indeterminate firms maximize their profits subject to the demand equation (3.10):

$$\max_{P_{t+s}(j), H_{t+s}(j)} E_t \left[ \sum_{s=0}^{\infty} \beta^s Q_{t|t+s} \frac{P_{t+s}(j)}{P_{t+s-1}(j)} \left( 1 - \Phi_p \left( \frac{P_{t+s}(j)}{P_{t+s-1}(j)} \right) \right) Y_{t+s}(j) - \frac{W_{t+s}}{P_{t+s}} H_{t+s}(j) \right]$$

$$Y_{t+s}(j) = \left( \frac{P_{t+s}(j)}{P_t} \right) Y_{t+s}$$

The solution for the optimal labor demand defines effective real marginal costs as

$$\frac{W_t}{P_t A_t} = \frac{P_t(j)}{P_t} \left( 1 - \Phi_p \left( \frac{P_t(j)}{P_{t-1}(j)} \right) \right) - \mu_t \equiv \frac{MC_t}{A_t} \quad (3.12)$$

The Lagrange multiplier  $\mu_t$  represents the marginal profit and equals the price corrected for the adjustment costs minus the marginal costs.

The decision for optimal price combined with (3.12) gives the Phillips Curve:

$$\begin{aligned} & \frac{1}{P_t} \left( 1 - \Phi_p \left( \frac{P_t(j)}{P_{t-1}(j)} \right) \right) Y_t(j) + \beta E_t \left[ Q_{t|t+1} \left( \frac{P_{t+1}(j)}{P_t(j)} \right)^2 \frac{Y_{t+1}(j)}{P_{t+1}} \Phi'_p \left( \frac{P_{t+1}(j)}{P_t(j)} \right) \right] \\ &= \left( \frac{1}{P_t} \left( 1 - \Phi_p \left( \frac{P_t(j)}{P_{t-1}(j)} \right) \right) - \frac{MC_t}{A_t P_t(j)} \right) \frac{Y_t(j)}{\lambda_{p,t}} + \frac{P_t(j)}{P_{t-1}(j) P_t} \Phi'_p \left( \frac{P_t(j)}{P_{t-1}(j)} \right) Y_t(j) \end{aligned}$$

The benefits of an increase in the price by one unit are given by the increase in revenue and the lower future expected adjustment costs (left hand side) are equal to its costs: a decrease in the revenue due to the decline in demand and the current price adjustment costs (right hand side). In a symmetric equilibrium all firms set equal prices and produce equal quantities. The aggregate Phillips curve is (3.13), where  $\pi_t = \frac{P_t}{P_{t-1}}$  denotes the gross inflation rate.

$$\pi_t \Phi'_p(\pi_t) + \frac{1 - \lambda_{p,t}}{\lambda_{p,t}} (1 - \Phi_p(\pi_t)) = \frac{MC_t}{A_t \lambda_{p,t}} + \beta E_t \left[ Q_{t|t+1} \frac{Y_{t+1}}{Y_t} \pi_{t+1} \Phi'_p(\pi_{t+1}) \right] \quad (3.13)$$

It is important to note that if the price adjustment costs are convex the Phillips curve is convex as well. Dolado et al. (2003, 2004) show that a convex aggregate supply curve leads to a nonlinear optimally-derived Taylor rule. More precisely, the central bank increases the interest rate more aggressively when inflation is above the target than when it is below the target. This asymmetry arises even when the objective function of the central bank is quadratic.<sup>15</sup>

**Nonlinear Taylor rule.** Monetary policy is assumed to obey the following nonlinear

---

<sup>15</sup>Even though the Dolado et al. (2004) paper does not find empirical evidence for the nonlinear Phillips curve in the US data after 1983, they do find asymmetries in the interest rate rule behavior with respect to inflation.



interest rate rule:

$$\begin{aligned} \ln R_t = & \rho_R \ln R_{t-1} + (1 - \rho_R) \ln R^* + (1 - \rho_R) \psi_y \left[ \ln \frac{Y_t}{Y_{t-1}^\gamma} + \frac{\exp\left(\psi_y^A \ln \frac{Y_t}{Y_{t-1}^\gamma}\right) - 1}{10\psi_y^A} \right] \\ & + (1 - \rho_R) \psi_\pi \left[ \ln \pi_t - \ln \pi^* + \frac{\exp\left(\psi_\pi^A (\ln \pi_t - \ln \pi^*)\right) - 1}{10\psi_\pi^A} \right] \epsilon_{R,t} \end{aligned} \quad (3.14)$$

Adjustments of the interest rate are determined by the sum of two linex functions. One depends on the deviations of inflation from its target. The other depends on the deviations of the output growth rate from the BGP. The degree and the sign of asymmetry are driven by the parameters of the linex functions. For example, if  $\psi_\pi = 1$  and  $\psi_\pi^A = 10$  then the interest rate decreases linearly for the values of the inflation below the target and increases exponentially for the values of the inflation above the target. Or if, for example,  $\psi_y = 1$  and  $\psi_y^A = -10$  then low output growth is highly undesirable. Whenever the output grows slower than on the BGP the interest rate decreases exponentially. However, when output is rising faster than on the BGP the interest rate only rises linearly (Appendix 3.7 presents the plot of this example). This formulation represents the idea that the central bank might have asymmetric preferences and introduces state dependent coefficients in the monetary policy rule. When asymmetry parameters approach zero, the monetary policy rule becomes identical to the standard Taylor rule. I estimate  $\psi_\pi$ ,  $\psi_y$ ,  $\psi_\pi^A$  and  $\psi_y^A$  from the data to determine the type and the magnitude of the asymmetries in the interest rate dynamics.

The formulation of the Taylor rule 3.14 has some limitations. For example, it implicitly assumes that the asymmetry in inflation deviations from its target and the asymmetry in the output growth rate movements are independent. Moreover, the interest rate smoothing parameter is constant. The model is, thus, unable to measure if the reaction to inflation is different in booms versus recessions or if the reaction to the output growth rate is affected by the current level of inflation. I also assume that the kink in the reaction to inflation appears at the inflation target (and in the reaction to the output growth rate at the target output growth rate). In reality it might be the case that the central bank changes its behavior only when inflation significantly deviates from the target so the kink must be higher or lower than the target inflation (the same is true for the output growth rate).

Note also that in a log-linearized version of the model, the interest rate reacts linearly to inflation and the output growth rate with the coefficients that are products of the  $\psi$ 's,  $\psi^A$ 's and other parameters. Scale coefficients  $\psi$ 's and asymmetry coefficients  $\psi^A$ 's cannot be separately identified. On the contrary, in the second-order approximate solution of the model

this equation contains separate terms for volatilities of the central bank target variables. The coefficient in front of the inflation volatility is determined solely by the asymmetry parameter  $\psi_\pi^A$ . The same is true for the volatility of the output growth rate. Consequently, I am able to identify the asymmetry parameters through the additional restrictions that the model imposes on the second order moments of the variables.

Finally, the fiscal authorities collect lump-sum taxes to finance exogenous government purchases expressed as a share of the output  $G_t = \frac{g_t-1}{g_t} Y_t$  where  $g_t$  follows an exogenous process:  $\ln g_t = (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \epsilon_{g,t}$ .

### 3.4 Estimation Results

The model was solved by the second order perturbation method around a non-stochastic steady state using the approach of Schmitt-Grohe and Uribe (2004).

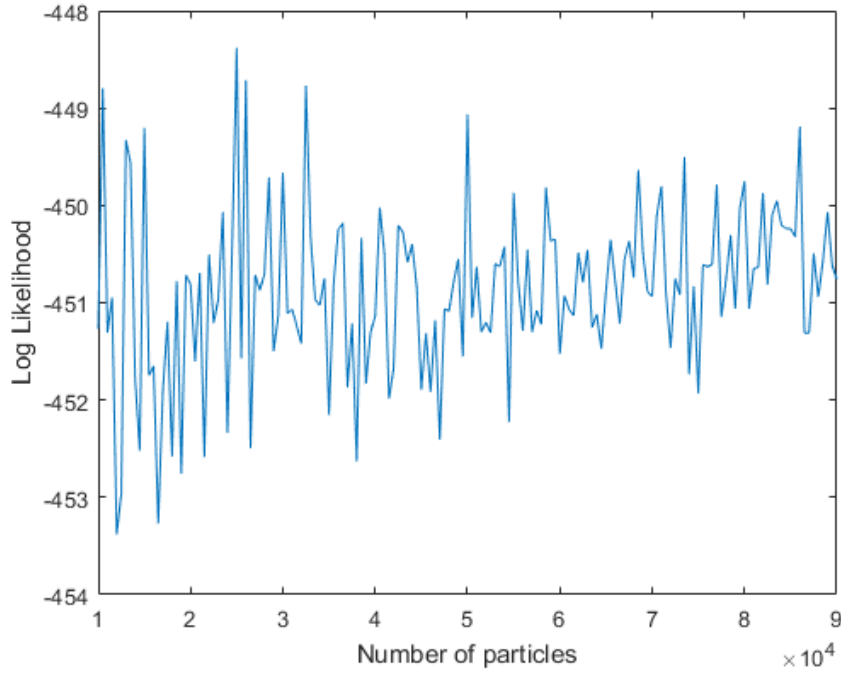
In the nonlinear context a standard Kalman filter cannot be applied to compute the likelihood. Instead the likelihood function is approximated using the sequential importance sampling algorithm - a particle filter.<sup>16</sup> Sequential Monte-Carlo methods can be applied to nonlinear models and models with non-Gaussian shocks. A brief summary of the estimation strategy is presented in the Appendix 3.7.

The measurement errors are assumed to have zero-mean and the variance is set to 10% of the variance of the corresponding observed variable. I used 100,000 particles for the likelihood approximation. A large number of particles is required for an efficient and smooth likelihood approximation. An upper limit is imposed by the computational time. The standard practice is to increase the number of particles sequentially until the variance of the point approximation (normally at the maximum of linearized likelihood) falls below a particular level. Figure 3.6 shows that at around 100,000 particles the approximation of the likelihood becomes more robust to the exact number of particles and the variance of the approximation declines.

Posterior distributions for the parameters were obtained by the Random Walk Metropolis Hastings with 85,000 draws and 50,000 burn-in draws. The scale parameter for the variance of the proposal density was adjusted such that the average acceptance rate was between 30 and 45%. The first-order approximation of the likelihood was maximized to obtain the mean and the variance for the proposal distribution. To a large extent I utilize the code of

---

<sup>16</sup>The general procedure is described in Doucet, de Freitas and Gordon (2001) or Andrieu, Doucet and Holenstein (2010). Macroeconomic applications are provided in Fernández-Villaverde and Rubio-Ramírez (2007) and in Herbst and Schorfheide (2015) among others.



**Figure 3.6:** Log-Likelihood at Prior Mode for Different Number of Particles

Frank Schorfheide.<sup>17</sup> The solved model is simulated by the pruning algorithm of Kim, Kim, Schaumburg and Sims (2003).

All priors and posteriors are summarized in the Table 3.2. Most of the priors are adopted from Aruoba et al. (2017) and based on the regressions on the pre-sample data. For the parameters that determine the overall reaction of the interest rate to the deviations of the inflation and the output growth I keep the standard gamma priors centered around conventional values. For the asymmetry parameters in the interest rate rule I choose uniform priors on the interval  $[-100, 100]$  to stay agnostic about the sign and the magnitude of the asymmetry. Figures (3.11) and (3.12) in the Appendix plot the recursive averages and marginal densities of the parameter estimates.<sup>18</sup>

The estimates of the standard parameters are broadly in line with the literature. Estimates for the price and wage adjustment costs imply that prices and wages are downward rigid. The results further indicate significant asymmetries in the interest rate rule. The coefficient  $\psi_\pi^A$  is small and positive. The point estimate is 7.2. This implies that the reaction of the

<sup>17</sup>Available at <https://web.sas.upenn.edu/schorf/publications/>.

<sup>18</sup>Note that the marginal densities for the last parameters that determine asymmetry -  $\psi_\pi^A$ ,  $\psi_y^A$ ,  $\psi_p$  and  $\psi_w$  - do not resemble normal distribution. This comes from the fact that their priors are flat. Thus, the shape of the posterior is determined by the shape of the nonlinear likelihood approximation from the particle filter.

**Table 3.2:** Posterior Estimates for DSGE Model Parameters

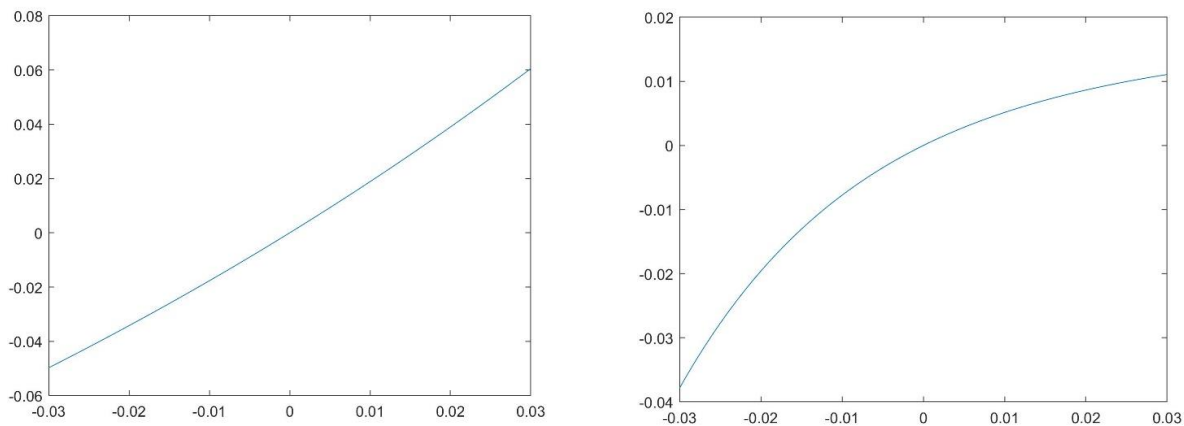
Parameter		Prior		Initial value	Posterior Mean	90% Interval
		Mean	Std			
$\tau$	Gamma	2.00	1.00	4	5.42	[4.68, 6.31]
$\nu$	Gamma	0.50	1.00	0.1	0.09	[0.08, 0.11]
$400\left(\frac{1}{\beta} - 1\right)$	Gamma	2.00	1.00	2	1.42	[1.06, 1.88]
$\pi^A$	Gamma	3.00	1.00	3	3.45	[2.89, 4.05]
$\gamma^A$	Gamma	2.00	1.50	2	2.18	[1.81, 2.56]
$\kappa(\varphi_p)$	Gamma	0.30	0.20	0.2	0.16	[0.09, 0.25]
$\psi_\pi$	Gamma	1.50	0.50	1.5	2.11	[1.74, 2.60]
$\psi_y$	Gamma	0.20	0.10	0.7	0.72	[0.41, 0.95]
$\varphi_w$	Gamma	15.0	7.50	12	13.7	[8.61, 20.5]
$\rho_r$	Beta	0.50	0.20	0.7	0.79	[0.73, 0.83]
$\rho_g$	Beta	0.80	0.10	0.85	0.96	[0.95, 0.97]
$\rho_z$	Beta	0.20	0.10	0.1	0.11	[0.03, 0.20]
$\rho_p$	Beta	0.60	0.20	0.85	0.92	[0.86, 0.96]
$100\sigma_g$	InvGamma	0.75	2.00	0.8	0.62	[0.48, 0.84]
$100\sigma_z$	InvGamma	0.75	2.00	0.5	0.51	[0.48, 0.57]
$100\sigma_p$	InvGamma	0.75	2.00	5	7.71	[6.01, 9.72]
$100\sigma_r$	InvGamma	0.20	2.00	0.2	0.15	[0.12, 0.17]
$\psi_w$	Uniform	-200	200	60	160.9	[147.8, 174.3]
$\psi_p$	Uniform	-300	300	165	121.7	[98.3, 131.49]
$\psi_\pi^A$	Uniform	-100	100	1	7.17	[0.85, 15.131]
$\psi_y^A$	Uniform	-100	100	1	-45.8	[-59.15, -32.6]

**Note:** For InvGamma distribution the shape and scale parameters are given instead of the mean and std.

central bank to inflation is linear when inflation is below the target and exponential when inflation is above the target. In other words, the central bank reacts more strongly to high inflation. For the output growth rate the asymmetry coefficient is negative and large with the point estimate -45.8. It means that the central bank decreases the interest rate more aggressively when the output growth is slow and increases the interest rate only moderately when the output growth is above the target.

Figure (3.7) plots the estimated reaction of the interest rate to inflation and the output

growth rate. For example, a 3% increase of inflation above its target is associated with a 6% increase of the interest rate while a 3% decrease of the inflation below the target results in a decrease of the interest rate by approximately 5%. When the the GDP growth is 3% above the BGP, interest rate is increased by approximately 1%. If the GDP grows 3% slower than on the BGP, the interest rate is decreased by about 3.8%.



**Figure 3.7:** Estimated Reaction of the Interest Rate to Inflation (left) and Output Growth (right)

In this context, one natural way to assess the accuracy of the model is to compare the marginal data density under the considered model and under an alternative one. Model comparison based on the ratio of the marginal data densities favors the model with the best one-step-ahead predictive performance<sup>19</sup> and, under some regularity conditions, the model that is asymptotically closest to the data generating process as measured by the Kullback-Leibler divergence (see Fernández-Villaverde and Rubio-Ramírez, 2004).

I compare two version of the model - one with the asymmetric interest rate rule and one with the standard linear Taylor rule. The standard Taylor rule was obtained by setting the asymmetry parameters in the linex function to zero ( $\psi_y^A = \psi_\pi^A = 0$ ). To compute the marginal density of the model with the standard rule I repeated the parameter estimation

<sup>19</sup>To see that note that the marginal data density is a product of one step ahead prediction errors  $p(Y_{1:T}|\theta) = \prod_{t=1}^T p(Y_t|Y_{0:t-1}, \theta)$ . An, Schorfheide (2007) provide more details.

procedure using the same priors. Posterior odds ratios are computed as:

$$\frac{P(M_1|Y)}{P(M_2|Y)} = \frac{p(M_1) * f(Y|M_1)}{p(M_2) * f(Y|M_2)} \quad (3.15)$$

$$f(Y|M) = \int \mathcal{L}(\theta|Y)p(\theta)d\theta \quad (3.16)$$

where  $p(\cdot)$ 's are prior probabilities of the models  $M_1$  and  $M_2$ , which I assume to be equal. The second term is a ratio of the marginal data densities  $f(Y|M)$ , each computed under the null hypothesis that the corresponding model is the true model.  $f(Y|M)$  is computed by Geweke's (1998) modified harmonic mean estimator.

**Table 3.3:** Posterior odds ratios (for different truncation probability)

Model	Log marginal data density		
	$\tau=0.1$	$\tau=0.5$	$\tau=0.9$
Linear rule	-519.8	-518.2	-517.6
Asymmetric rule	-465.1	-465.0	-465.4
Log Bayes factor	54.7	53.2	52.3

**Note:**  $\tau$  is a truncation probability parameter for Geweke (1998) harmonic mean estimator

Table (3.3) shows the results for the posterior odds ratio test. One can conclude that the data speak in favor of the model with asymmetric reaction of the central bank. The model with asymmetric Taylor rule is  $\exp(54.7)$  times more likely.

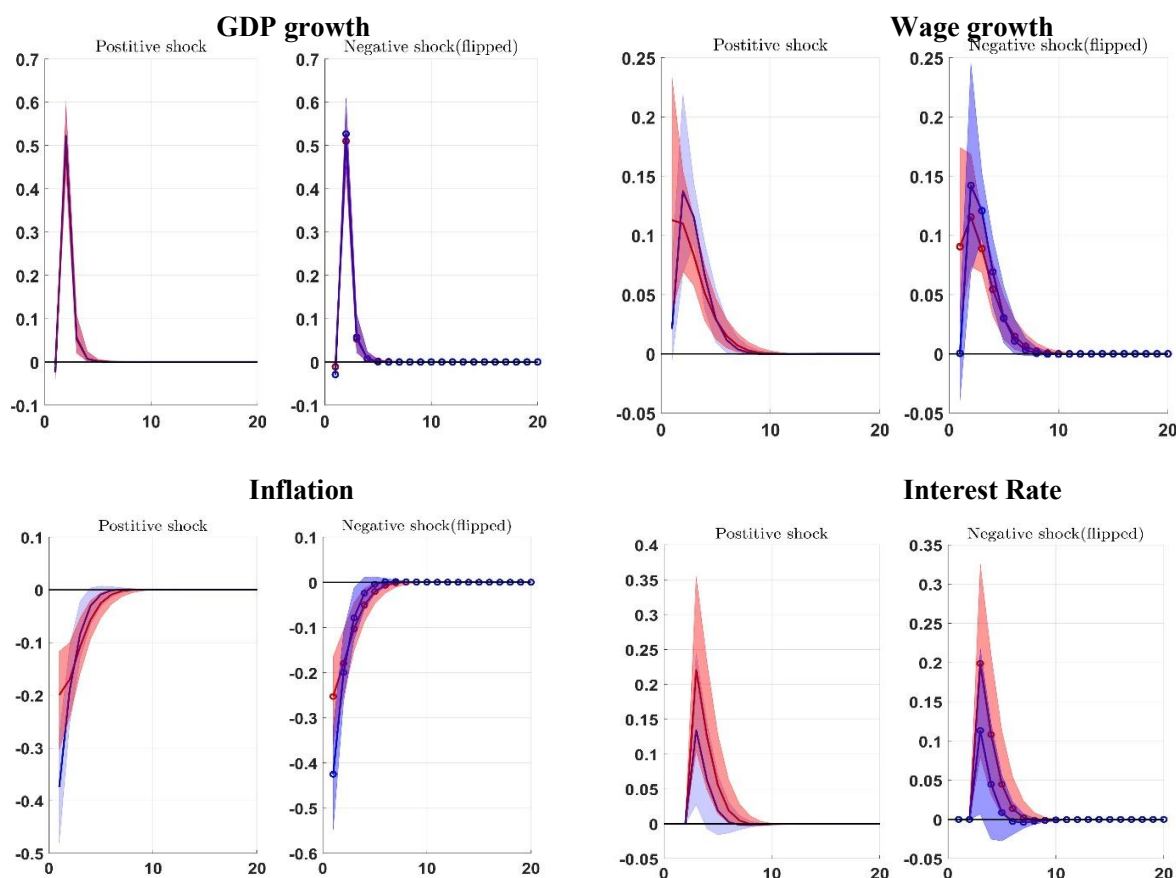
I now turn to the analysis of how the asymmetric reaction function of the central bank affects economic dynamics. The propagation of the shocks in the economy can be analyzed by impulse response functions. In a nonlinear model these functions become state and shock dependent. Koop, Pesaran and Potter (1996) proposed to use generalized impulse response functions (GIRF) to characterize the dynamics in the nonlinear systems. GIRFs are computed as (3.17).

$$GIRF_{t+k}(v_i, w_t) = E[y_{t+k}|\epsilon_{t+1} = v_i, w_t] - E[y_{t+k}|w_t] \quad (3.17)$$

The response of the economy to a shock that hits in period  $t$  depends on the current state of the economy ( $w_t$ ), on the size and the sign of the shock ( $v_i$ ) as well as on the whole history of the shocks that occurred since  $t$ .

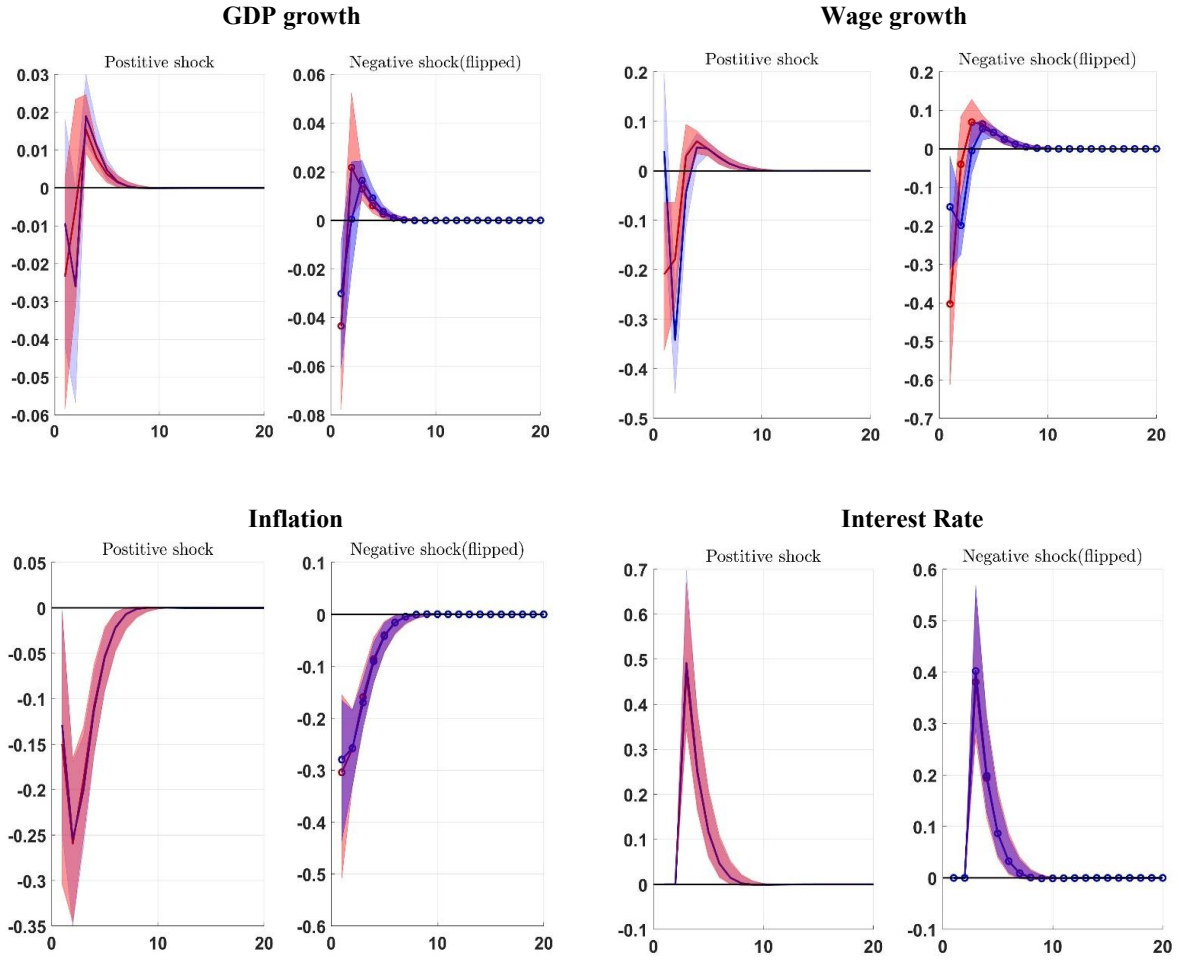
I plot GIRFs for the technology and for the monetary policy shock each of 1 standard deviation. The shock can hit either in a recession regime or in a boom regime. I define a recession as a period of the posterior simulation with the lowest GDP growth and a boom as

a period with the highest GDP growth (similar to Amisano, Tristani, 2010). I set to zero all future shocks but one of the interest.



**Figure 3.8:** Generalized Impulse Response Functions for 1sd Technology Shock: Expansion (red) vs. Recession (blue)

Figures (3.8) and (3.9) present the resulting GIRFs. The main insight from the figure is that the 90% credible sets of the responses largely overlap. One thus cannot reject the hypothesis that the reaction of the economy is symmetric in the current state and the sign of the shock. The largest discrepancy in the mean response can, perhaps, be seen in the reaction of inflation and the interest rate to a positive technology shock. If a positive shock to the output growth rate hits in the recession, inflation rate declines by about 0.35%. When the positive shock occurs in a boom inflation goes down by only 0.2%. Monetary policy reacts to it by increasing the interest rate. However, in a recession the central bank has less incentives to raise the interest rate since the decline of the inflation is larger. For the negative technology shocks these asymmetries are rather insignificant.



**Figure 3.9:** Generalized Impulse Response Functions for 1sd Monetary Policy Shock: Expansion (red) vs. Recession (blue)

The GIRFs indicate that the model dynamics remain highly symmetric across states and different signs of the shock despite the presence on nonlinearities in the inflation, wage and interest rate dynamics. Increasing the size of the shocks to 2 or 3 standard deviations does not affect this conclusion. The computation of GIRFs by averaging out the simulated path of the future shock would be an interesting exercise to explore in the future research. In the next section, I show that dynamics of the interest rate in the model also fails to pass QAR posterior predictive checks despite the nonlinear Taylor rule.

### 3.5 Posterior Predictive Checks

Aruoba et al. (2017) propose to evaluate a nonlinear DSGE model by a nonlinear version of the posterior predictive checks.

The main idea of the method is to represent the solution of a DSGE model as a nonlinear



state space model - quadratic autoregressive model (QAR):

$$y_t = \phi_0 + \phi_1(y_{t-1} - \phi_0) + \phi_2 s_{t-1}^2 + (1 + \gamma s_{t-1})\sigma u_t \quad (3.18)$$

$$s_t = \phi_1 s_{t-1} + \sigma u_t, \quad u_t \text{ iid} \sim N(0, 1) \quad (3.19)$$

For the current application the most important coefficient is  $\phi_2$  that measures asymmetries of the data.

**The procedure.** QAR predictive checks are conducted as follows.

1. First the  $\theta_{QAR} = [\phi_0, \phi_1, \phi_2, \gamma, \sigma^2]$  is estimated for the actual data series. The likelihood of this model is estimated iteratively by assuming that  $y_0$  and  $s_0$  are known and solving the system for  $u_t$  and  $s_t$  at every step recursively. Posteriors for the  $\theta_{QAR}$  is then obtained by combining the likelihood with priors. For  $\phi_0, \phi_2, \gamma$  Aruoba et al. (2017) assume normal priors. The mean of  $\phi_0$  is set to the mean of the corresponding data series in the pre-sample period.  $\phi_1$  is distributed according to the truncated normal distribution with the center at the AR(1) coefficient estimated from pre-sample data.  $\sigma$  follows the inverse gamma distribution centered around the standard deviation of the residuals from the AR(1) model estimated on the pre-sample data. The distribution for initial values  $y_0$  and  $s_0$  are obtained by simulating the system for  $t=-T:t$ , where  $T$  is set to 20. The random walk Metropolis-Hastings procedure was used to generate draws from the posterior distribution for  $\theta_{QAR}$ .
2. On the second step the same  $\theta_{QAR}$  are obtained for the data simulated from the DSGE model. At every iteration a draw  $\theta$  is taken from the posterior of the DSGE estimation. The model is simulated with this parameter vector for 1000 periods. Further the  $\theta_{QAR}$  is estimated on this simulated data as described above. 2000 draws are taken from the posterior for the QAR parameters and the median is saved. This procedure is repeated 1000 times. The resulting time series of 1000 medians represent an empirical posterior distribution for  $\theta_{QAR}$ .

Table (3.4) compares the QAR coefficients for the interest rate in the actual data, in the model with the standard Taylor rule and in the model with the asymmetric rule. The  $\phi_2$  coefficient (the corresponding row is marked in bold) is large and negative in the interest rate data. However, in both versions of the model it is very close to zero. The model with the asymmetric interest rate rule is unable to generate a realistic degree of asymmetry in the interest rate.

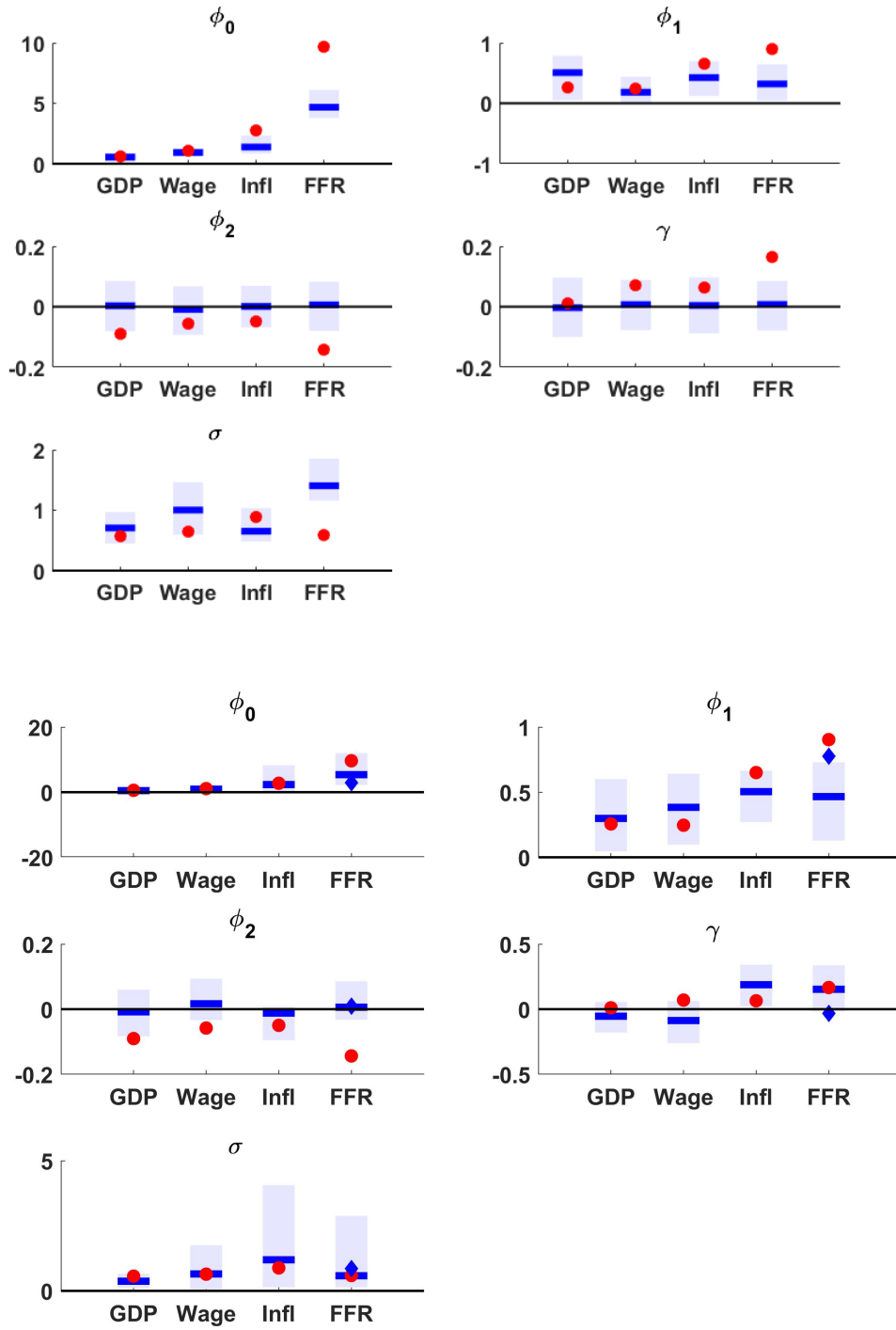
**Table 3.4:** QAR posterior predictive checks

$\theta_{QAR}$	DSGE Posterior Means		
	Data	Standard TR	Asymmetric TR
$\phi_0$	9.80	6.59	7.18
$\phi_1$	0.91	0.57	0.56
$\phi_2$	<b>-0.16</b>	<b>0.00</b>	<b>0.00</b>
$\gamma$	0.08	0.00	0.00
$\sigma$	0.22	2.2	2.12
$\sigma$	0.25	0.41	0.40

**Note:** Asymmetric Taylor rule is defined by equation (3.14). Standard Taylor rule is obtained from the asymmetric one by setting  $\psi_\pi^A = \psi_y^A = 0$ . Numbers are rounded up to the second digit.

Figure (3.10) replicates the original analysis of the Aruoba et al. (2017) and presents the results for the model with an asymmetric Taylor rule. One can conclude that the largest discrepancy between the data and the nonlinear DSGE is the asymmetry coefficient for the Federal funds rate. This result holds regardless of the asymmetries in the interest rate rule.<sup>20</sup> Finally, I test the hypothesis that the asymmetries in the interest rate data are driven by the asymmetric behavior of the inflation and the output growth rate themselves. If the central bank target variables behave asymmetrically then even under a linear policy rule the dynamics of the interest rate would be asymmetric. To test this hypothesis I computed the dynamics of the interest rate implied by the estimated nonlinear Taylor rule. However, I took the actual data for inflation and GDP growth instead of simulations. In other words, I ask what would happen if the model could reproduce the dynamics of inflation and GDP growth perfectly. I then run posterior predictive checks on this artificial interest rate data. The results are indicated as blue diamonds on the bottom panel plots of Figure (3.10). The  $\phi_1$  coefficient of the artificial data is closer to the coefficient for the actual data. However, the credible set for the coefficient of the asymmetry  $\phi_2$  is still centered around zero and the point estimate  $\phi_2 = -0.0257$  is still too small. In the actual data this coefficient is -0.16. This result implies that asymmetric targets per se would not resolve the problem of the implausibly small asymmetries in the interest rate dynamics in the model.

<sup>20</sup>These results are robust to setting the wage and price asymmetries to zero. Moreover, similar QAR coefficients are preserved when the asymmetric Taylor rule is formulated with a logistic function or tanh function.



**Figure 3.10:** QAR Coefficients: Standard Taylor Rule (upper panel) and Asymmetric Taylor Rule (lower panel).

*Notes:* Red dots indicate the actual data, bars are the 90% credible sets for the simulated series.

### 3.6 Conclusion

This paper evaluates asymmetries in the interest rate dynamics. It contributes to the literature in the following directions. First, it estimates the degree of asymmetry in the monetary policy rule within a nonlinear DSGE model. According to my results, the US central bank takes stronger actions during periods of low output growth than during periods of high output growth and during periods of high inflation than during periods of low inflation. Posterior odds tests favor the model with the asymmetric Taylor rule.

Second, this paper contributes to the discussion on the ability of a nonlinear DSGE model to reproduce a plausible asymmetric behavior of the interest rate. Aruoba et al. (2017) run QAR posterior predictive checks and show that a standard DSGE with a linear Taylor rule fails to generate the dynamics of the interest rate with realistic higher order moments. In the current study I show that the asymmetric Taylor rule does not improve the results of the posterior predictive checks. The model is still unable to generate a plausible degree of the asymmetry in the interest rate dynamics.

Additionally, the analysis in this paper sheds light on whether the asymmetric monetary rule affects the propagation of the shocks in the model. I plot the generalized impulse response functions (GIRF) after a technology shock and after a monetary policy shock for 1) different states of the economy and 2) for different signs and magnitudes of the shocks. According to the results these GIRFs are statistically indistinguishable. The propagation of the shocks in the current model is highly symmetric.

The overall conclusion of the paper is that asymmetry in interest rate dynamics cannot be explained solely by the asymmetric reaction of the monetary policy. Other mechanisms might contribute to the observed behavior of the interest rates. Lindé and Trabandt (2018), for example, argue that Kimball aggregation is essential to reproduce the skewness of macroeconomic variables in a nonlinear DSGE. A different specification of the asymmetries in the interest rate rule might help to bring the model closer to the data. Another explanation could be that asymmetries in the interest rate movements arise from the non-systematic reaction of the central bank, in other words, that the monetary policy shocks are asymmetric. This would be in accordance with the Sims and Zha (2006) paper. They find little evidence on changes in the coefficients of the monetary policy rule in the US but strong evidence on time varying volatility of the structural shocks.

Several policy-relevant conclusions emerge from my results. Many central banks cut the interest rates to very low levels during the Great recession. In the course of the crises

interest rates hit the zero-lower-bound in many economies and conventional monetary became ineffective. According to this paper one of the reasons for why interest rates reached zero might be the overreaction of the central bank to economic crises due to recession avoidance preferences. If this is true then the suggestions to increase the interest rate target to avoid hitting the zero-lower-bound in the future might be less effective than it is generally considered. Interest rates would reach the zero-lower-bound constraint faster than estimated.

Another implication relates to the effectiveness of the monetary policy. If rational agents have learned the asymmetric behavior of the central bank then they would expect the same reaction in the future and incorporate this fact in their decisions. For example, after a negative demand shock firms would expect a large reduction of the interest rate before they start to re-optimize their investment strategy. If monetary policy deviates from the asymmetric rule that it has been following historically and reduces the interest rate by less, then the effect of the expansionary monetary intervention might be mitigated.

The conclusions of this paper open two avenues for future research. First, this study leaves an open question about which features of a nonlinear DSGE model are essential for a satisfactory model performance as evaluated by the QAR posterior predictive checks. Second, the macroeconomic consequences of the asymmetric monetary policy behavior that was identified in this paper deserve further exploration.

### 3.7 Appendix

#### Nonlinear Estimation Summary

Consider a standard problem of computing a posterior of the DSGE model parameters  $p(\theta|Y_{1:T})$ :

$$\begin{aligned} p(\theta|Y_{1:T}) &= \frac{p(Y|\theta)p(\theta)}{f(Y)} \\ p(Y_{1:T}|\theta) &= \prod_{t=1}^T p(Y_t|Y_{0:t-1}, \theta) \\ f(Y) &= \int p(Y|\theta)p(\theta)d\theta \end{aligned}$$

A standard linear state space model can be written as:

$$\begin{aligned} y_t &= \Psi(\theta)s_t + u_t, \quad u_t \sim N(0, \Sigma_u) \\ s_t &= \Phi(\theta)s_{t-1} + \epsilon_t \quad \epsilon_t \sim N(0, \Sigma_\epsilon) \end{aligned}$$

Non-linear/Non-Gaussian state space model takes a more general form:

$$\begin{aligned} y_t &= \Psi(s_t, \theta) + u_t, \quad u_t \sim F_u(\dots, \theta) \\ s_t &= \Phi(s_{t-1}, \epsilon_t, \theta) \quad \epsilon_t \sim F_\epsilon(\dots; \theta) \\ &\downarrow \\ \ln \mathcal{L}(Y_{1:T}|\theta) &= \sum_t^T \ln [p(Y_t|Y_{1:t-1}, \theta)] \end{aligned}$$

Since the likelihood function of the model is non-Gaussian, it is not enough to track solely the mean and the variance of this function. Instead one must approximate the whole distribution.

I use a Bootstrap Particle Filter which can be summarized as follows.

#### BPF Algorithm:

##### Step 1: Initialization

draw initial particles  $s_0^1, s_0^2, \dots, s_0^M \sim p(s_0)$ . Set initial weights  $W_0^j=1$ ,  $j=1 \dots M$ .

For  $t=1, \dots, T$ :

##### Step 2: Forecasting

propagate the states (particles) and compute the likelihood at each particle

$$\begin{aligned}\tilde{s}_t^j &= \Phi(s_{t-1}^j, \epsilon_t^j | \theta) \quad \epsilon_t^j \sim F_\epsilon(\cdot) \\ p(y_t | \tilde{s}_t^j, \theta) &= \int \delta(y_t - \Psi(\tilde{s}_t^j, u_t, \theta)) p_u(u_t | \theta) d(u_t) \\ &= p_u(y_t - \Psi(\tilde{s}_t^j, \theta) | \theta)\end{aligned}$$

### Step 3: Updating

Update the weights of the particles.  $\{W^j, s^j\}$  then define a discrete distribution.

$$\tilde{W}_t^j = \frac{p(y_t | \tilde{s}_t^j, \theta) W_{t-1}^j}{\frac{1}{M} \sum_j p(y_t | \tilde{s}_t^j, \theta) W_{t-1}^j}$$

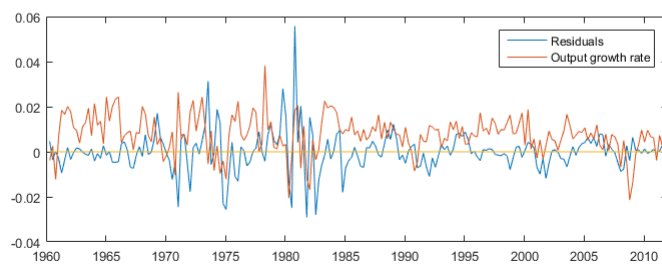
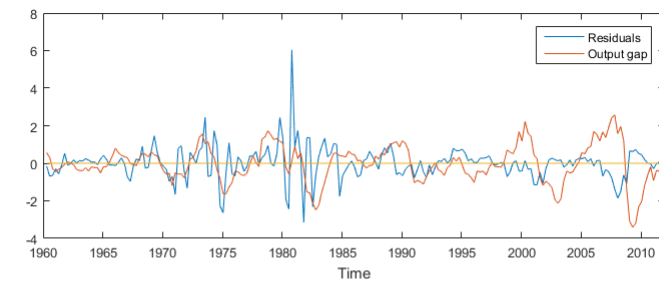
### Step 4: Resampling

If an effective sample size falls below a threshold resample from  $\{\tilde{W}_t^j, \tilde{s}_t^j\}$

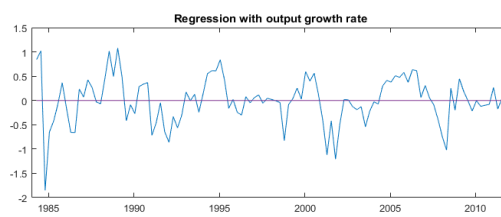
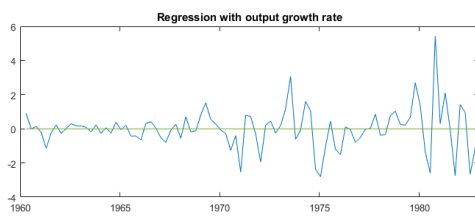
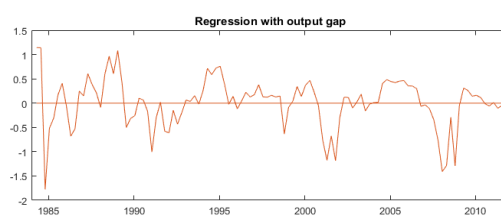
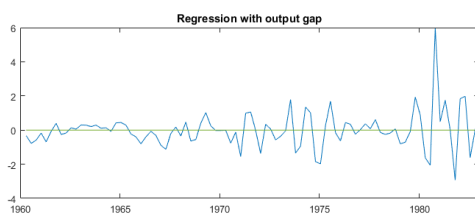
**The likelihood is given by**

$$\ln \mathcal{L}(Y_{1:T} | \theta) = \sum_t^T \ln \left( \frac{1}{M} \sum_j^M p(y_t | \tilde{s}_t^j, \theta) W_{t-1}^j \right)$$

## Taylor Residuals, 1960-2012

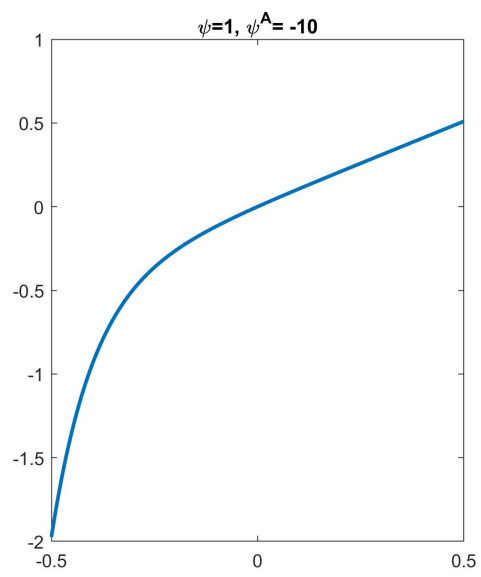
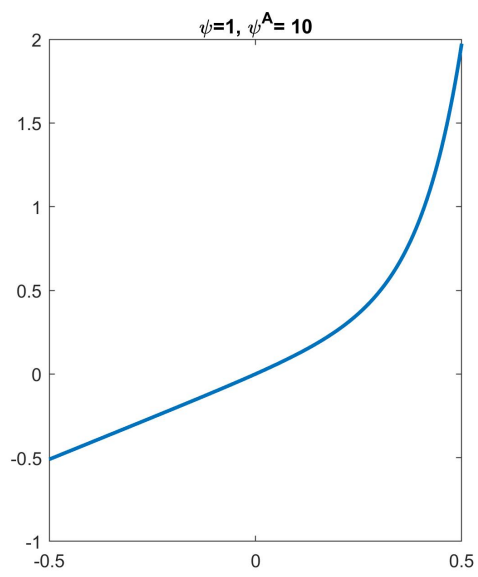


## Taylor Residuals for Sub-periods, 1960-1983 and 1984-2012





## Linex Function for Interest Rate Rule (Example)



## Estimation Results

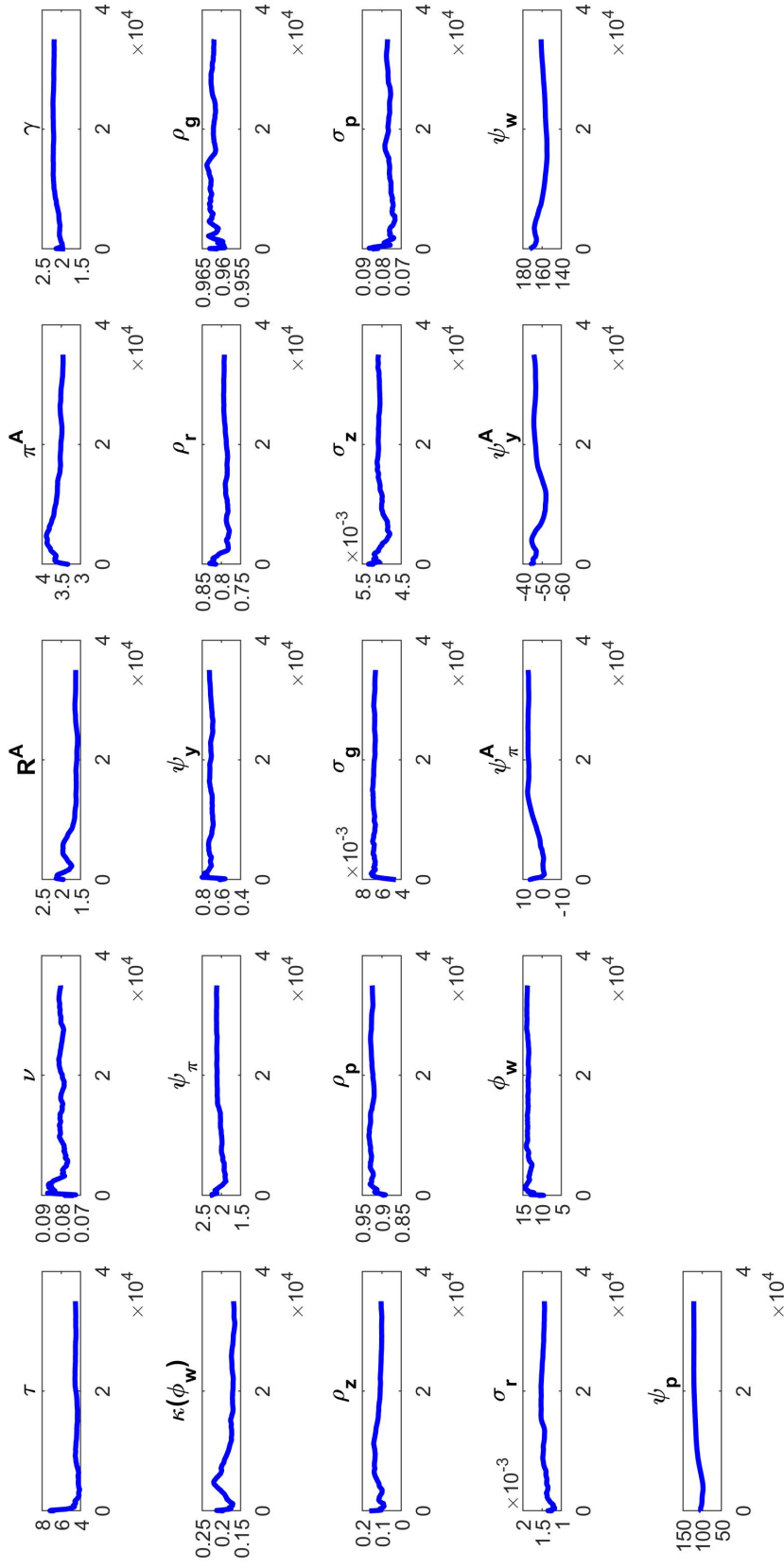


Figure 3.11: Recursive Averages

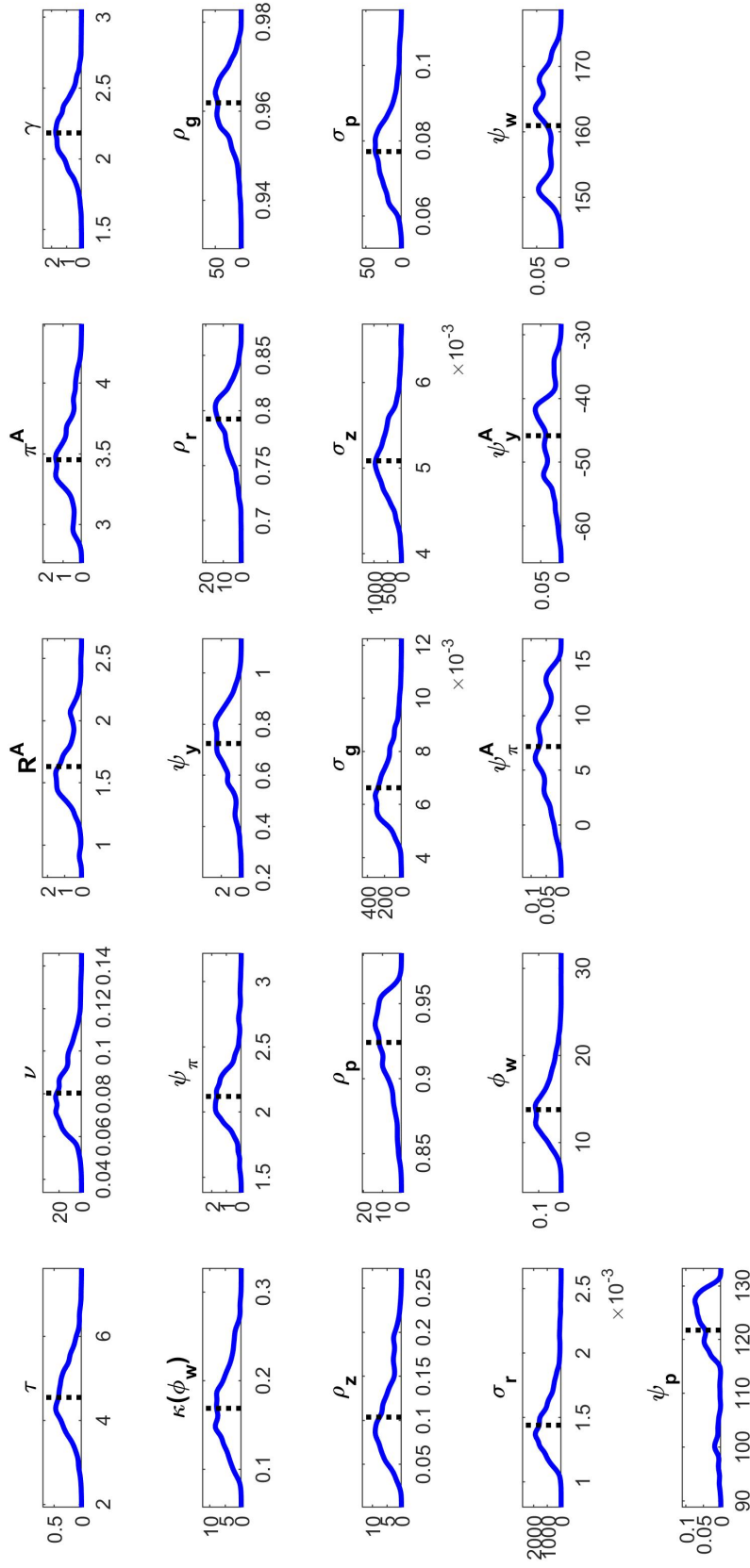


Figure 3.12: Marginal Densities





## 4 New Challenges for Monetary Policy: Digital Currencies and Currency Competition

### Abstract

In this paper I use a workhorse of monetary economics - the Lagos and Wright (2005) search theoretic model augmented with privately issued currencies as in Fernández-Villaverde and Sanches (2016). First, I extend the model by linear costs that associated with the private currency circulation. I show that in contrast to Fernández-Villaverde and Sanches (2016) digital currency competition 1) does not deliver price stability and 2) puts a downward pressure on the public currency inflation only when the costs of private currency circulation are sufficiently low. Second, I build a model of the economy with a single generally accepted blockchain-based currency. In the spirit of the search and matching literature I use a matching function to model how miners process transaction requests. I show that in a blockchain-based monetary system money demand features a precautionary motive which is absent in the standard Lagos-Wright model. Due to this precautionary money demand monetary equilibrium is stable for some calibrations. I use the developed model to study how the equilibrium inflation depends on the blockchain parameters such as mining costs and rewards.

**Keywords:** Currency competition, Cryptocurrency, Miners, Inflation, Blockchain

**JEL classification:** E40, E41, E42, E50, E58

### 4.1 Introduction

More than 2100 different blockchain-based digital currencies with a total market capitalization of about \$100 billion were in circulation in January, 2019.<sup>21</sup> Over 300,000 transactions are conducted daily with cryptocurrencies and at least 1,000 off-line locations accept cryptocurrency as means of payment.<sup>22</sup>

The appearance of blockchain-based currencies poses many normative and positive questions for monetary economics (Fernández-Villaverde and Sanches, 2017). These questions require a theoretical model that takes into account the most important characteristics of the blockchain

---

<sup>21</sup> [www.coinmarketcap.com](http://www.coinmarketcap.com).

<sup>22</sup> The numbers are for Bitcoin, a cryptocurrency with the largest capitalization. See <https://www.coindesk.com/surge-in-real-locations-accepting-bitcoin/> and <https://www.coindesk.com/surge-in-real-locations-accepting-bitcoin/>.

operation. To the best of my knowledge, no such monetary model has yet been developed.

In this essay I study two important questions that arise due to the emergence of the private digital currencies. First, what is the outcome of the competition between private digital currencies and the fiat currency of the central bank. Second, how properties of the blockchain protocol affect monetary equilibrium. The novelty of this study is that I explicitly take into account the most important features of the blockchain protocol.

I argue that cryptocurrencies have two important characteristics that distinguish them from the other forms of privately supplied money.<sup>23</sup> The first is the absence of currency issuers (except ICOs - initial currency offerings). Newly created currency units are injected into the network as a reward for the execution of financial transactions. The processing of financial transactions is called mining and the agents who conduct this activity are called miners. I want to stress that mining is conceptually different from currency issuance. Mining requires computational time as input and costs energy.<sup>24</sup> Mining costs and rewards determine the equilibrium supply of the mining service and as a byproduct affect the evolution of the money supply.

The second important feature of the blockchain is that processing of financial transactions takes time. This is a necessary requirement for the operation of the blockchain which results from the double-spending problem and is not an artefact of technological constraints. Hence, not all transaction that were requested are processed within a period (or ever). Money market participants take a probabilistic nature of the transaction execution into account when signing trade contracts. They can add a fee to a transaction request to facilitate the execution. The circulation of the blockchain-based currency features externalities. Miners compete with each other to be the first who process a particular transaction since only the first miner receives the reward for this transaction. Hence every miner is successful only with a particular probability. An increase in the total number of miners negatively affects the probability of a miner's success. On the other hand, an increase in the number of the transaction requests positively affects the miners' revenue through an increase in transaction fees. In a similar way the probability that a particular transaction will be processed increases with the number of miners but decreases in the transaction demand.

This essay builds on Fernández-Villaverde and Sanches (2016)<sup>25</sup> - henceforth FVS - who

---

<sup>23</sup>An example from the history is the free banking era in the US in 1837-1863 or Scotland until 1865 (see Gordon, 1985).

<sup>24</sup>For example, an operation of blockchains with a proof-of-work is associated with significant energy costs (See, O'Dwyer and Malone, 2014).

<sup>25</sup>This paper cites Fernández-Villaverde and Sanches (2016) and was written independently from the later

augment the Lagos-Wright model<sup>26</sup> with privately supplied currencies. In their framework the currency issuers follow a commitment rule. This assumption represents the fact that supply of cryptocurrencies is determined by a computer algorithm and cannot be changed in a discretionary fashion. FVS show that the equilibrium with currency competition is efficient (in the sense that it delivers price stability) but is unstable.

The first part of this essay modifies the FVS model by acknowledging that transactions involving cryptocurrency are associated with linear operational costs (similar approach can be found in Marimon et al., 2003). I show that in this case the outcome of the currency competition is inefficient. The monetary equilibrium with privately supplied currencies does not deliver price stability. Instead, a positive inflation rate is required to compensate currency suppliers.

I then analyze competition between private and public currencies. I assume that government supplies public money at no costs to maximize its seigniorage. I show that currency competition works as a disciplinary device and imposes an upper bound on the equilibrium inflation rate that the government can sustain. This result is standard in the literature on currency competition. However, this conclusion only holds if the costs of private currency circulation do not exceed a particular threshold level. Above this threshold, inflation in the equilibrium with competitive currency provision is higher than the seigniorage maximizing inflation. Consequently, currency competition plays no role for monetary policy.

In the second part of this essay I study the economy with a single generally accepted blockchain-based currency. I develop a monetary model that can be used for the formal analysis of the blockchain-based monetary system. Processing of financial transactions in this model is described by a matching function in the spirit of the search and matching literature. The idea of a matching function originates in labor economics<sup>27</sup>. It is widely used in macro models to describe market imperfections: for example, labor market frictions<sup>28</sup> or the credit market mismatch<sup>29</sup>. The matching function naturally reproduces externalities associated with mining and allows to determine the probability of a miner success as well as the probability of a transaction processing. Both probabilities become functions of the

---

revision in October, 2017. The results of the 2017 version and this paper are on par.

<sup>26</sup>One of the advantages of the Lagos-Wright (2005) framework is an endogenous money demand. Alternative frameworks are transaction costs, money-in-the-utility, cash-in-advance constraint or an OLG model (see Sargent, 2009).

<sup>27</sup>Pissarides(2000), Mortensen and Pissarides (1994, 1999).

<sup>28</sup>See Gali (2010) for a review.

<sup>29</sup>For example, as den Haan et al., 2003, or Wasmer and Weil (2004).



number of transaction requests relative to the number of active miners. The probability of transaction processing corresponds to the exogenous probability of trade in the Lagos-Wright model. However, it becomes endogenous in the current model.

I derive monetary equilibrium of the model in terms of the rate of return on money. Since the model explicitly includes blockchain characteristics - mining costs and rewards - I am able to analyze their effects on the equilibrium money return in a comparative statics exercise. A higher reward to miners accelerates money growth and reduces the equilibrium return on money. Higher costs of mining, on the contrary, put upward pressure on the equilibrium return on money.

I also show that agents in the economy hold an excess amount of currency units and request an excess amount of transactions. Since agents know that every transaction request is processed only with a particular probability they submit more of them out of a "precautionary motive". As a consequence, the money demand function is hump-shaped: it increases for small values of the return on money and starts to decrease after some point. If the monetary equilibrium corresponds to the declining part of the money demand then the equilibrium is stable. These results distinguish my model from models of traditional centralized monetary system with a central bank.

Many ideas of this paper can be traced back to the literature on currency competition and the time consistency of monetary policy - Klein (1974) and Taub (1985) among others. Since at least von Hayek (1976) monetary economics tries to answer whether competitively supplied money is more efficient than a monopoly of the central bank on the money market. These papers, however, do not model costs of money circulation. Marimon et al. (2012) do discuss the costs of currency provision and show that the equilibrium inflation level is increasing in the cost parameter. However, they assume that private currency issuers are subject to the lack of commitment in the same manner as the monetary authority. Marimon et al. (2003) is closely related to the analysis of this paper. They analyze a costly provision of currency substitutes by banks in a model with a cash-in-advance constraint. Similar to their paper I show that currency competition forces the government to set a lower inflation rate than it would prefer (for small enough cost parameter). However, this paper differs from Marimon et al. (2003) in that I treat currency units solely as means of payment in anonymous trade. No real asset and no form of credit can be used as a payment in such an environment. On contrary, Marimon et al. (2003) consider bank deposits as currency substitutes. Currency issuers - banks - have an access to interest bearing government bonds which means that the

return on money and the return on the real bonds must be equalized. As a result in their model currency competition drives the upper bound of inflation to a negative level as dictated by the Friedman rule.

Other papers that analyse currency competition in a search-theoretic framework include Berentsen (2006), Cavalcanti et al. (1999) and Cavalcanti and Wallace (1999). These papers, however, do not address transaction costs or other features of modern cryptocurrencies.

A growing amount of the literature addresses the economic nature and the novelty of blockchains.<sup>30</sup> Most of the papers, however, focus on case studies or a general discussion of the blockchain protocol. Theoretical literature on private currency provision and currency competition is extensive.<sup>31</sup> However, the existing models lack many of the fundamental features of the blockchain operation and therefore can only partially be applied to study blockchain-based monetary systems.

The rest of the paper is organized as follows. Section 2 describes the model of currency competition. It addresses both the equilibrium with purely private currency provision and the competition between the private and the central bank money. Section 3 presents a monetary model of the blockchain operation and shows how the properties of the mining process affect monetary equilibrium. Section 4 concludes.

## 4.2 A Model of Digital Currency Competition

This section presents a monetary model of the competition between digital currencies and the central bank public currency. First I discuss the money demand function from Lagos-Wright (2005). Second I describe how miners process monetary transactions and introduce mining costs. I then analyze currency competition among private currencies and between private currencies and the central bank currency.

### 4.2.1 Money Demand

Money demand in the model is identical to Fernandez-Villaverde and Sanches (2016) which is a replication of Lagos and Wright (2005) search theoretic approach. I briefly sketch the structure of the economy and redirect a reader to the above mentioned papers for detailed derivations.

---

<sup>30</sup> Berentsen and Schaer (2018), Swan (2015), Peters et al. (2016).

<sup>31</sup> For example, Cavalcanti, Erosa, and Temzelides (1999, 2005), Cavalcanti and Wallace (1999), Williamson (1999), Berentsen (2006), Marimon et al. (2003, 2012) and more recently, Fernández-Villaverde and Sanches (2017).

The economy is populated by a continuum of buyers of measure 1, a continuum of sellers of measure 1 and infinitely many miners. Every period  $t$  consists of two sub-periods - a centralized market in which every agent can produce and exchange a centralized good  $x$  and a decentralized market in which only sellers produce and only buyers buy and consume a decentralized good  $q$ . Production function of the sellers is linear and takes labor as the only input.

During the centralized market phase buyers and sellers are endowed with a fixed amount of labor and can produce and consume centralized goods. On the decentralized market buyers meet sellers with a probability  $\sigma$ . Every pair negotiates on a deal according to which a seller produces and sells a particular amount  $q$  of goods to the buyer. Negotiation is achieved by a "take-it-or-leave-it" offer. This type of bargaining eliminates additional inefficiencies arising from the bargaining process itself.

The buyer and seller in a pair meet once and never again. This means that it is not possible to use credit in the trading process. Instead, the agents need to use a storable generally accepted medium of exchange - money.<sup>32</sup> Need for money arises endogenously from the fundamental structure of the economy. No other real assets can be used as means of payment.

The buyer's problem reads as

$$\begin{cases} W_t^b(m_t^b) = \max_{x_t^b, \hat{m}_t^b} [x_t^b + \sigma (u(q_t(\hat{m}_t^b, \hat{m}_t^s)) + \beta W_{t+1}^b(\hat{m}_t^b - p_t(\hat{m}_t^b, \hat{m}_t^s))) + (1 - \sigma)\beta W_{t+1}^b(\hat{m}_t^b)] \\ \text{s.t. } \phi_t \hat{m}_t^b + x_t^b = \phi_t m_t^b \end{cases}$$

$$W_t^b(m_t^b) = \phi_t m_t^b + \max_{\hat{m}_t^b} [-\phi_t \hat{m}_t^b + \sigma (u(q_t(\hat{m}_t^b, \hat{m}_t^s)) + \beta W_{t+1}^b(\hat{m}_t^b - p_t(\hat{m}_t^b, \hat{m}_t^s))) + (1 - \sigma)\beta W_{t+1}^b(\hat{m}_t^b)]$$

Here,  $W_t^b(m_t^b)$  is the buyer's utility function which depends on the buyer's money holdings during the centralized market stage denoted as  $m_t^b$ ,  $x_t^b$  is the buyer's consumption on the centralized market,  $u(q)$  is the utility from consuming the good  $q$  on the decentralized market,  $p_t$  denotes the payment that the buyer has to pay to the seller on the decentralized market,  $\phi$  is the real value of money unit and  $\beta$  is a discount factor. Variables with hat  $\hat{m}_t^s$  and  $\hat{m}_t^b$

---

<sup>32</sup>In the Lagos and Wright (2005) paper every agent is assumed to produce a unique good and to have preferences over the goods of others. Money is required to overcome the double coincidence problem. The double coincidence means that each party in the pair wants the good of counter-party and barter is possible. An exogenous distinction between buyers and sellers that I borrowed from Fernández-Villaverde and Sanches (2016) does not affect the main result.

denote the seller's and the buyer's money holdings during the decentralized market. The utility function has standard properties:  $u(0) = 0, u'(0) = \infty, u'(\cdot) > 0, u''(\cdot) < 0$ .

A seller solves an analogous problem:

$$\begin{cases} W_t^s(m_t^s) = \max_{x_t^s, \hat{m}_t^s} [x_t^s + \sigma [-w(q_t(\hat{m}_t^b, \hat{m}_t^s)) + \beta W_{t+1}^s(\hat{m}_t^s + p_t(\hat{m}_t^b, \hat{m}_t^s))] + (1 - \sigma)\beta W_{t+1}^s(\hat{m}_t^s)] \\ \text{s.t. } \phi_t \hat{m}_t^s + x_t^s = \phi_t m_t^s \end{cases}$$

where  $w(n_t^s)$  is a disutility from labor and the production function is linear  $q_t = n_t^s$ . Standard assumptions on the disutility function hold:  $w(0) = 0, w'(\cdot) > 0, w''(\cdot) > 0$ .

Negotiation on the decentralized market is described by the following problem. The buyer maximizes his utility subject to a participation constraint for the seller (PC) that ensures that the seller does not make losses and to a liability constraint (LC) that guarantees that the buyer cannot pay more money than he has. Under such a framework money holdings of the seller play no role for the bargaining outcome. The bargaining can be characterized in terms of  $q$  and  $p$  which are both functions of  $\hat{m}_t^b$ .

$$\max_{q_t, p_t} [u(q_t) - \beta \phi_{t+1} p_t] \quad (4.1)$$

$$\text{s.t. } -w(q_t) + \beta \phi_{t+1} p_t \geq 0 \quad \text{PC} \quad (4.2)$$

$$p_t \leq \hat{m}_t^b \quad \text{LC} \quad (4.3)$$

A standard result in the contract theory is that the PC is always binding. The objective function (4.1) can be written as  $u(q_t) - w(q_t)$ . Optimal  $q^*$  is then defined from  $u'(q^*) = w'(q^*)$ . If the buyer has enough money to compensate the seller for a production of  $q^*$  then  $q^*$  is produced. The optimal payment can be derived from the PC. If the buyer's money holdings are insufficient he simply pays everything he has and receives an amount  $q$  that the seller is willing to produce for the compensation  $\hat{m}_t^b$ .

$$q_t = \begin{cases} q^* & \text{if } \phi_{t+1} \hat{m}_t^b \geq \beta^{-1} w(q^*) \\ w^{-1}(\beta \phi_{t+1} \hat{m}_t^b) & \text{if } \phi_{t+1} \hat{m}_t^b < \beta^{-1} w(q^*) \end{cases}$$

$$\phi_{t+1} p_t = \begin{cases} \beta^{-1} w(q^*) & \text{if } \phi_{t+1} \hat{m}_t^b \geq \beta^{-1} w(q^*) \\ \phi_{t+1} \hat{m}_t^b & \text{if } \phi_{t+1} \hat{m}_t^b < \beta^{-1} w(q^*) \end{cases}$$

FOC of the buyer's problem is:  $\frac{dW^b(\hat{m}_t^b)}{d\hat{m}_t^b} \leq 0$  and  $= 0$  if  $\hat{m}_t^b > 0$ . Thus, for any currency  $i$  that is held in equilibrium

$$\begin{cases} -\phi_t^i + \beta \phi_{t+1}^i = 0 & \text{if } \phi_{t+1} \hat{m}_t^b \geq \beta^{-1} w(q^*) \\ \sigma u'(q_t) q'_{m_t^i} + \beta(1 - \sigma) \phi_{t+1}^i - \phi_t^i = 0 & \text{if } \phi_{t+1} \hat{m}_t^b < \beta^{-1} w(q^*) \end{cases}$$

Where  $q$  is a function of the total real money holdings  $q(\hat{m}_t^b) = w^{-1}(\beta\phi_{t+1}\hat{m}_t^b)$ . Taking a derivative of the inverse function  $q'_{\hat{m}_t^i} = \beta\phi_{t+1}^i \frac{1}{w'(q)}$ . The second equation can be written as  $\frac{\phi_t^i}{\beta\phi_{t+1}^i} = 1 + \sigma \frac{u'(q)}{w'(q)} - \sigma$ .

A solution exists only with  $\frac{\phi_{t+1}}{\phi_t} \leq \beta^{-1}$ . Indeed, if the return to money  $\gamma_{t+1} \equiv \frac{\phi_{t+1}}{\phi_t}$  is larger than the inverse of the time preference parameter than agents would want to hold an infinite amount of money units. When  $\frac{\phi_{t+1}}{\phi_t} \leq \beta^{-1}$  holding money is costly. Since the bargaining and the trade on the decentralized market depend solely on buyers' money holdings only buyers decide to hold currencies and  $m_t^s = 0$ . From now on I will use  $\hat{m}_t = \hat{m}_t^b$  which denotes the total money demand in the economy.

One can show that the objective function of a buyer (4.2.1) is concave and is strictly decreasing at  $\beta^{-1}w(q^*)$  if  $\frac{\phi_{t+1}}{\phi_t} \leq \beta^{-1}$  (see Appendix 4.5). It means that the buyer's problem has a unique solution when  $\phi_{t+1}\hat{m}_t < \beta^{-1}w(q^*)$ . In a limiting case when  $\phi_t = \beta\phi_{t+1}$  the derivative of the objective function is zero at  $\beta^{-1}w(q^*)$ . In this situation I assume a limiting solution  $\hat{m}_t = \beta^{-1}w(q^*)$  as in Lagos and Wright (2005).

Money demand can be written as:

$$\frac{\phi_t^i}{\beta\phi_{t+1}^i} = 1 + \frac{\sigma}{\beta\phi_{t+1}^i} [u'(q) - w'(q)] q'_{m_t^i} \quad (4.4)$$

$$\frac{\phi_t^i}{\beta\phi_{t+1}^i} = 1 + \frac{\sigma}{\beta\phi_{t+1}^i} [u'(q) - w'(q)] \beta\phi_{t+1}^i \frac{1}{w'(q)} \quad (4.5)$$

$$\frac{\phi_t^i}{\beta\phi_{t+1}^i} = 1 + \sigma \frac{u'(q)}{w'(q)} - \sigma \quad (4.6)$$

#### 4.2.2 Private Currencies

Infinitely many private agents, called miners, provide private cryptocurrencies which are perfect substitutes for each other. Each miner maintains a blockchain of a cryptocurrency  $i$ . Let us assume that  $N$  blockchains are in operation.

Given a vector of money holdings  $m_t = (m_t^1, \dots, m_t^N)'$  the buyer solves a utility maximization problem by deciding how much money to spend on consumption  $x_t^b$  and how much to save for the decentralized market phase  $\hat{m}_t = (\hat{m}_t^1, \dots, \hat{m}_t^N)'$ . The real value of a unit of the  $i$  the currency in terms of a consumption good is  $\phi_t^i$  and the total real value of money /holdings is  $\sum_{i=1}^N \phi_t^i m_t^{bi} = \phi_t m_t^b$  where  $\phi_t = (\phi_t^1, \dots, \phi_t^N)$ .

Let us denote the return to money as  $\gamma_{t+1} \equiv \frac{\phi_{t+1}}{\phi_t}$ . Equation (4.6) can be rewritten as (4.7). The money demand in the economy is defined by two functions: the production on the decentralized market as a function of the return on money  $q(\gamma_{t+1})$  (4.7) and the total real value of money holdings as a function of the level of production  $q$  (4.8). Note that since (4.7)

holds for any currency  $i$ , the return on money must be equalized among currencies that are held in equilibrium.

$$\sigma \frac{u'(q_t)}{w'(q_t)} + 1 - \sigma = \frac{\phi_t}{\beta \phi_{t+1}} = \frac{1}{\beta \gamma_{t+1}} \quad (4.7)$$

$$\phi_{t+1} \hat{m}_t = \beta^{-1} w(q_t) \quad (4.8)$$

The real money demand can be expressed as  $z(\gamma_{t+1}) \equiv \phi_t \hat{m}_t = \sum_{i=0}^N \frac{\phi_{t+1}^i}{\phi_{t+1}^i} \phi_t^i \hat{m}_t^i = \phi_{t+1} \hat{m}_t \frac{1}{\gamma_{t+1}}$ .

$$z(\gamma_{t+1}) = \frac{\beta^{-1} w(q(\gamma_{t+1}))}{\gamma_{t+1}} \quad (4.9)$$

### 4.2.3 Equilibrium with Private Money

Private cryptocurrencies are provided by rational private agents - miners - who maximize the linear utility function that depends on the the stream of consumption  $x^{M33}$  subject to linear operational costs with parameter  $\psi$ .

Miners' behavior is determined by an algorithm of a blockchain protocol which is a public knowledge. All agents, therefore, know the future evolution of the money supply as in a perfect foresight set up.

$$\max_{\{M_t^i\}} \sum_{t=0}^{\infty} \beta^t x_t^{Mi} \quad (4.10)$$

$$\text{s.t. } x_t^{Mi} = \phi_t (M_t^i - M_{t-1}^i) - \psi \phi_t^i M_t^i \quad (4.11)$$

Here  $0 < \psi < 1$  is a fraction of the real money stock that is spent on the blockchain operation. Linear operational costs are similar to financial intermediation costs in Marimon et al. (2003). In case of the Bitcoin blockchain, for example, these are the costs of electricity that is used for transaction verification (proof-of-work). These costs represent the inefficiency of the private money provision technology. I use a linear formulation for simplicity but it can be easily substituted by any functional form.

Is it plausible to assume that the total number of processed transactions and, therefore, the energy costs are increasing with the total real money stock? Or in other words, should we multiply  $\psi$  by the money stock and not with the change in money stock  $M_t^i - M_{t-1}^i$ . The key feature of the blockchain that lies behind this assumption is that mining does not correspond to money issuance directly. Mining is an activity of processing transactions. When existing currency units are transferred back and forth from one account to another each transaction

---

<sup>33</sup>In the real world miners indirectly affect the money supply by optimally choosing the intensity of mining activity for a particular level of mining costs.

has to be processed. Even if the money supply stays constant so  $M_t^i - M_{t-1}^i = 0$  a circulation of the existing money units still consumes energy. Additionally, it is logical to assume that a higher purchasing power of a currency unit leads to a higher number of transactions with this currency and, hence, higher total costs of processing these transactions.

The market for each cryptocurrency features free entry for miners, which implies zero profit of miners in equilibrium.

$$\sum_{t=0}^{\infty} \beta^t [(1 - \psi) \phi_t^i M_t^i - \phi_t^i M_{t-1}^i] = 0 \quad \forall i \quad (4.12)$$

Consumption level is non-negative every period  $(1 - \psi) \phi_t^i M_t^i - \phi_t^i M_{t-1}^i \geq 0 \quad \forall t$ . Consequently,

$$(1 - \psi) M_t^i - M_{t-1}^i = 0 \quad \forall t \text{ if } \phi_t^i > 0 \quad (4.13)$$

Linear costs of mining require a positive money supply growth at a rate  $\frac{\psi}{1-\psi}$  such that miners are compensated by positive seigniorage.<sup>34</sup> Note that the money growth rate cannot be negative.

In equilibrium the total real money supply  $\sum_i^N \phi_t^i M_t^i$  must be equal to the real money demand  $z(\gamma_{t+1})$ . The share of every currency in the buyer's portfolios is indeterminate but since for all valued currencies the supply grows at the same rate and the return on money must be equal, we can write

$$M_t^i = \frac{1}{1 - \psi} M_{t-1}^i = \left( \frac{1}{1 - \psi} \right)^t M_0^i$$

$$\phi_{t+1}^i = \phi_t^i \gamma_{t+1} = \phi_0^i \prod_{\tau=0}^t \gamma_{\tau+1}$$

Let us assume that the equilibrium is symmetric, that is  $M_0^i = M_0^j$  and  $\phi_0^i = \phi_0^j \quad \forall i, j$ . Then every valued currency has an equal weight in the buyer's portfolios in every period. Combining (4.13) and (4.9) the equilibrium path is described by

$$(1 - \psi) z(\gamma_{t+1}) - \gamma_t z(\gamma_t) = 0 \quad (4.14)$$

which defines  $\gamma_{t+1}(\gamma_t)$ . This equation has two stationary points:  $\gamma = 0$  and  $\gamma = 1 - \psi$ . Monetary equilibrium with  $\gamma = 1 - \psi$  implies a decreasing real price of money or in other words a positive inflation rate.

**Proposition 1:** In a monetary equilibrium  $\gamma = 1 - \psi < 1$  which means that the real price of money is declining and the inflation level is strictly positive.

---

<sup>34</sup>If the cost parameter were different for different currencies than the growth rates of money supply would also differ.

Currency competition cannot deliver price stability. As long as the costs associated with the operation of private currency protocols are nonzero, a positive seigniorage is required to compensate miners for their costs. This result is different from FVS who show that a competitive currency supply brings about a constant price equilibrium.

FVS, moreover, show that the monetary equilibrium is unstable. This result also holds in the current set up.

**Proposition 2:** A monetary equilibrium with a positive return on money is unstable. Non-monetary equilibrium with a zero return on money is stable.

*Proof:* As long as money demand function satisfies  $z'(\gamma) > 0$  the equilibrium with a positive value of money is unstable. If initial conditions are such that  $\gamma_0 < 1 - \psi$  then from equation (4.14)  $\gamma_t$  declines for  $t \geq 1$ . When  $\gamma_0 > 1 - \psi$  the equilibrium path is explosive. Formally, using the implicit function theorem and the assumption that  $z'(\gamma) > 0$ :

$$\begin{aligned} \frac{d\gamma_{t+1}}{d\gamma_t} &= \frac{z(\gamma_t) + \gamma_t z'(\gamma_t)}{(1 - \psi)z'(\gamma_{t+1})} \\ \left. \frac{d\gamma_{t+1}}{d\gamma_t} \right|_{\gamma_t = \gamma_{t+1} = 1 - \psi} &= \frac{z(1 - \psi) + (1 - \psi)z'(1 - \psi)}{(1 - \psi)z'(1 - \psi)} > 1 \end{aligned}$$

thus,  $\gamma_t \rightarrow 0$  if  $\gamma_0 < 1 - \psi$  and  $\gamma_t \rightarrow \infty$  if  $\gamma_0 > 1 - \psi$ . As  $\gamma$  converges to zero, real money demand converges to zero and the economy converges to a non-monetary equilibrium. Since  $\gamma_t$  is the same for all valued currencies, money demand goes to zero for all of them.<sup>35</sup> Appendix 4.3 plots the function  $\gamma_{t+1}(\gamma_t)$ .

To sum up, the equilibrium with private currency provision is neither efficient (it does not deliver a price stability) nor stable. Clearly, the costs of currency circulation is harmful for a welfare. Inefficient (smaller than 1) return on money leads to an inefficiently low money demand and a low amount of trade on the decentralized market.

**Proposition 3:** In a monetary equilibrium with  $\gamma = 1 - \psi < 1$  the volume of trade in the decentralized market  $q$  is decreasing in  $\psi$ .

*Proof:* Compare a level of trade in a non-zero costs equilibrium ( $q$ ) and in a zero costs equilibrium ( $q^0$ ). (4.7) must hold for any equilibrium.

$$\sigma \frac{u'(q)}{w'(q)} + 1 - \sigma = \frac{1}{\beta(1 - \psi)} \quad \text{and} \quad \sigma \frac{u'(q^0)}{w'(q^0)} + 1 - \sigma = \frac{1}{\beta}$$

substracting the second equation from the first gives:

$$\sigma \left( \frac{u'(q)}{w'(q)} - \frac{u'(q^0)}{w'(q^0)} \right) = \frac{1}{\beta(1 - \psi)} - \frac{1}{\beta} > 0$$

---

<sup>35</sup>This proof is identical to FVS (2016).



The last condition means that either  $u'(q) > u'(q^0)$  or  $w'(q) < w'(q^0)$  or both. Since  $u'(q)$  is decreasing and  $w'(q)$  is increasing it implies that  $q < q^0$ . The larger the parameter  $\psi$  the larger is the difference between  $q$  and  $q^0$ . Costs of private currency circulation are, thus, welfare-reducing.

The first-best production level  $q^*$  which is given by  $u'(q^*) = w'(q^*)$  is achieved when  $\gamma = \beta^{-1}$ . From the Fisher equation  $1 + i_t = (1 + r)(1 + \pi_t)$  and the fact that  $\gamma_t = \frac{1}{1 + \pi_t}$ , this implies that  $i = 0$ . To see that formally rewrite (4.7) as

$$\begin{aligned}\sigma \frac{u'(q_t)}{w'(q_t)} + 1 - \sigma &= \frac{1 + \pi}{\beta} \\ \sigma \frac{u'(q_t)}{w'(q_t)} &= 1 + i_t - 1 + \sigma \\ \frac{u'(q_t)}{w'(q_t)} &= 1 + \frac{i_t}{\sigma}\end{aligned}\tag{4.15}$$

$q = q^*$  if  $i_t = 0$ .

The first-best solution for the return on money corresponds to the Friedman (1969) rule. Friedman argued that the opportunity costs of money holding - the nominal interest rate - should be equal to zero and the economy should be saturated with money. In this case there will be no inefficiency in the real sector arising from monetary scarcity. Private currency competition in the current set up never achieves the Friedman rule because it requires a negative seigniorage. In the current model the seigniorage from a currency creation equals to the consumption of a miner which cannot be negative.

#### 4.2.4 Equilibrium with Public Money

In this section I introduce a government that provides public money taking potential private currency issuers into account. In this section I show that 1) currency competition imposes an upper limit on the level of inflation for a seigniorage-maximizing government if the costs of private currency circulation are small enough and that 2) this upper bound is positive.

I assume that the government finances a sequence of transfers  $G_t$  from its seigniorage. A path for the public money supply  $M^G$  is determined by a no-Ponzi games condition together with an intertemporal budget constraint

$$G_t = \phi_t^G(M_t^G - M_{t-1}^G)\tag{4.16}$$

If additionally the government has an ability to impose lump-sum taxes and  $G_t = \phi_t(M_t^G - M_{t-1}^G) + \tau_t$  then the government can constantly shrink money supply and achieve an efficient allocation in the decentralized market. Indeed, if the benevolent government follows the

Friedman rule by setting the return on public money  $\gamma^G$  to be  $\beta^{-1}$  the production in the decentralized market is equal to  $q^*$ .

Following the literature, I assume that the government decides on the money stock  $M_t$  which in turn defines the government transfers  $G_t$ . The government follows a constant money growth rule, e.i.  $M_t^G = (1+\omega)M_{t-1}^G$  and wants to maximize the stream of transfers  $\sum_{t=0}^{\infty} \beta^t G_t$  by choosing  $\omega$  rather than to rely on taxation. I denote the real value of public money as  $m_t^G = \phi_t^G M_t^G$  and return on public money as  $\gamma^G$ . As for private currencies for public currency  $\gamma_t^G = \frac{1}{1+\pi_t^G} = \frac{1}{1+\omega}$ . The government takes the money demand  $z(\gamma_t)$  as given.

The budget constraint (4.16) can be rewritten as

$$\sum_{t=0}^{\infty} \beta^t G_t = -\gamma_0^G m_{-1}^G + \sum_{t=0}^{\infty} \beta^{t+1} \left( \frac{1}{\beta} - \gamma_{t+1}^G \right) m(\gamma_{t+1}^G) \quad (4.17)$$

Since the government maintains a monopoly over the supply of public money it faces a time-inconsistency problem. Namely,  $-\gamma_t^G m_{t-1}^G$  equals to zero only in the period  $t = 0$  and is positive afterwards. The government, thus, has an incentive to relaunch the public currency every period. Under commitment the government decides on constant  $\gamma^G$  (see Marimon et al., 2003) to maximize a period by period seigniorage  $f(\gamma_{t+1}) = \left( \frac{1}{\beta} - \gamma_{t+1} \right) z(\gamma_{t+1})$ .

The inflationary tax  $\left( \frac{1}{\beta} - \gamma_{t+1}^G \right)$  is decreasing in  $\gamma^G$  (increasing in the inflation rate) whereas the money demand is increasing in the return on money (decreasing in the inflation rate). For the standard utility and production functions the seigniorage  $f(\gamma_{t+1})$  has a unique maximum  $\gamma^{G*}$ . Competition from the private sector, however, forces the government to set the return on currency to at least  $\gamma = 1 - \psi$ . Thus, the solution is restricted by  $1 + \omega \leq \frac{1}{1-\psi}$ .

If  $f'(\gamma)$  is negative at  $\gamma = 1 - \psi$  then the government would like to set  $\gamma^G$  lower than than it is forced by currency competition. This happens if

$$\begin{aligned} f'(\gamma) &= -z(\gamma) + \left( \frac{1}{\beta} - \gamma \right) z'(\gamma) < 0 \text{ or} \\ f'(\gamma) &= z(\gamma) \left[ -1 + \left( \frac{1}{\beta} - \gamma \right) \frac{z'(\gamma)}{z(\gamma)} \right] < 0 \end{aligned}$$

We can concentrate on the expression in the brackets since  $z(\gamma) \geq 0 \forall \gamma$ . I assume the same functional forms as in FVS:  $u(q) = \frac{q^{1-\eta}}{1-\eta}$ ,  $0 < \eta < 1$  and  $w(q) = \frac{q^{1+\alpha}}{1+\alpha}$ ,  $\alpha \geq 0$ . Given these specifications  $f'(\gamma)$  is negative if

$$\begin{aligned} &\left( \frac{1}{\beta} - \gamma \right) \left[ \frac{1-\eta}{\eta+\alpha} \frac{1}{\gamma} + \frac{1+\alpha}{\eta+\alpha} \frac{(1-\sigma)\beta}{1-(1-\sigma)\beta\gamma} \right] - 1 < 0 \\ &\text{substituting } 1-\psi \text{ for } \gamma \\ &\left( \frac{1}{\beta} - 1 + \psi \right) \left[ \frac{1-\eta}{\eta+\alpha} \frac{1}{1-\psi} + \frac{1+\alpha}{\eta+\alpha} \frac{(1-\sigma)\beta}{1-(1-\sigma)\beta(1-\psi)} \right] - 1 < 0 \end{aligned} \quad (4.18)$$

In general the sign of (4.18) depends on the parametrization. For example, Lagos and Wright (2005) set  $\beta = 0.997$  which corresponds to a 3% annual real interest rate,  $\alpha = 0$  which corresponds to a linear disutility from labor,  $\eta = 0.16$  and  $\sigma = 0.5$ . With these parameters a threshold level for  $\psi$  is 0.12. If a private currency provision is associated with costs that are higher than 0.12 then currency competition plays no role. If the costs of private money maintenance are lower than 12% per money unit then currency competition imposes a lower bound for the return on money (or equally an upper bound on the inflation level) that the government can sustain. The government can not set the return on public money lower than  $1 - \psi = 0.88$  (or inflation higher than 13%). Appendix 4.5 presents the threshold values for  $\psi$  for different calibrations.

**Proposition 4:** With  $u(q) = \frac{q^{1-\eta}}{1-\eta}$ ,  $0 < \eta < 1$  and  $w(q) = \frac{q^{1+\alpha}}{1+\alpha}$ ,  $\alpha \geq 0$  currency competition sets an upper bound for a sustainable inflation level only if  $\psi$  is smaller than the threshold level defined by (4.18).

If  $\psi$  goes to zero the outcome converges to a zero inflation equilibrium. However, as long as operational costs of private currencies are positive, the upper bound on the inflation is also positive. The government can collect a positive seigniorage even under a threat of competing private currencies.

To gauge a realistic value for the  $\psi$  parameter one can look at statistics for cryptocurrencies. For example, blockchain.info indicates for Bitcoin that the average mining costs per money unit is below 2%.<sup>36</sup> It means that if gross inflation level  $1 + \pi$  in some hypothetical country goes above  $1/(1 - \psi) \approx 1.02$  the currency holders would prefer Bitcoin over the fiat currency (abstracting from risk and trust concerns). From bitinfocharts one can compute the same number for other currencies. For etherum the cost parameter would be less than 0.5%.<sup>37</sup>

Finally, I discuss the situation when the government cannot commit to its policy. Imagine that the government follows a commitment rule until some period  $\tau$  and discretionary decides on  $\gamma_t$  every period for  $t > \tau$ . Marimon et al. (2003) show that the government will continue to follow the path of the commitment rule if there is a positive value attached to such policy - positive seigniorage. In the model of Marimon et al. (2003) private currency competition can drive inflation into a negative region. Private currency providers would bear a negative inflation rate because they have an access to other forms of income. With negative inflation

---

<sup>36</sup><https://blockchain.info/charts/cost-per-transaction-percent>

<sup>37</sup>Computed as reward in the last 24 hours over the amount of coins sent in the last 24 hours. See <https://bitinfocharts.com/ethereum/>

the government deviates from the commitment path and over-issues the public currency. As a result the government money becomes valueless and is driven out of circulation. In contrast, in the current framework inflation in the private currencies is always positive due to the presence of costs. Even with a lack of commitment private currencies never drive the public currency out of circulation.

The model predicts that high inflation in the public currency gives an incentive to the households to substitute private currencies for the public fiat money. This process indeed can be observed in some countries with hyperinflation or low trust in the national authorities. For example, Jack and Suri (2011) show that by the end of 2009 about 65% of households in Kenya were using cell phone currency M-Pesa for money transfers. M-Pesa was introduced in Kenya in 2007 by mobile provider Safaricom<sup>38</sup> and was represented by mobile phone minutes that could be send for large distances at extremely low costs with a help of a cell phone technology.

Another example would be Ecuador which introduced government digital cash controlled by the central bank. The willingness to accept the digital cash might stem from the hyperinflation episodes in Ecuador fiat currency until 2000.<sup>39</sup>

The decision to use private money instead of the legal tender crucially depends on the inflation level as well as on the trust in local government. Moreover, it matters whether a private currency is widely accepted as a mean of payment and whether it has a stable exchange rate. All these issues are not addressed in this paper.

### 4.3 A Monetary Model of a Blockchain

This section develops a monetary model with a single generally-accepted digital currency that circulates according to the blockchain protocol. The blockchain operation is modeled by a matching function analogously to labor search and matching literature. The formulation of the money demand is very similar to the Lagos-Wright (2005) approach except that the probability of trade on the decentralized market is endogenous in the current context and is derived from the matching function.

---

<sup>38</sup>Safaricom was redesigned to a public-private partnership with 50% owned by Vodafon and 25% by Kenya Treasury department.

<sup>39</sup><https://www.economist.com/blogs/americasview/2014/09/electronic-money-ecuador>

#### 4.3.1 Matching in the Money Market

The economy is populated by an infinity of miners who can decide to process financial transactions or stay inactive. Active miners randomly choose transaction requests, verify them and execute (which technically means add them to the blockchain). The total number of processed transactions  $T_t$  is specified as a matching function

$$T(d_t, CPU_t) = \bar{T} d_t^\eta CPU_t^{1-\eta} \quad (4.19)$$

where transaction demand  $d_t$  is the number of submitted transaction requests,  $CPU_t$  is the number of computer processing units that were employed for mining.  $0 < \eta < 1$  determines the elasticity of substitution between computational units and pending transaction requests.  $\bar{T}$  is a scale parameter.

According to the matching function both miners and transaction requests are needed to "create" new processed transactions on the blockchain. For example, if  $d_t = CPU_t$  then each miner "meets" a transaction request and  $T_t = d_t = CPU_t$ . If  $CPU_t < d_t$  then  $CPU_t < T_t < d_t$ . The framework accounts for the fact that some transactions remain pending when the transaction demand is too high. The other way around, if  $CPU_t > d_t$  then we have  $d_t < T_t < CPU_t$ . Some miners use their CPUs and spend electricity on mining but do not receive a reward. No transactions are processed when the demand for transactions is zero or when there are no miners in operation.

Several clarifications are in order. First, in reality miners process blocks of transactions. For simplicity I assume that one block contains only one transactions with only one currency unit. I also assume that one miner owns one computer processing unit so  $CPU_t$  is equal to the number of miners on the money market.

Second, from the technical point of view miners can mine empty blocks and still receive a reward. The matching function rules this out. I argue that mining of empty blocks is not sustainable in equilibrium. If no transactions are processed and added to the blockchain the currency can not be effectively used as a mean of exchange. The value of such currency is zero in the long run and the reward is valueless for the miners.

Third, the matching function states that an increase in  $CPU_t$  raises  $T_t$  which is not necessarily true in reality. Many of the blockchains periodically adjust the level of mining difficulty (and hence the mining costs) depending on the number of miners. These adjustments stabilize the number of transactions processed per unit of time. If, for example,  $CPU_t$  increases and, as a result,  $T_t$  rises then after a particular time period the difficulty is increased

and the number of verified transaction goes down again.<sup>40</sup> For such cryptocurrencies this model can be applied to characterize a short run equilibrium.

Forth, my specification implies that a higher transaction demand makes it easier for miners to "meet" transaction requests. This is not true for real blockchains. Miners can always find a pending transaction at no costs. However, higher transaction demand raises the fees that agents attach to their transaction requests and thus positively affects the profit of miners. I do not introduce fees explicitly. Instead, I assume that a positive effect of higher  $d_t$  comes from an increase in the probability of a miner's success.

Following the standard approach in the search and matching literature I characterize the money market in terms of market tightness  $\theta = \frac{d_t}{CPU_t}$  which is the number of transaction requests per miner.

The matching function naturally introduces externalities into the model. The probability of transaction confirmation  $\sigma(\theta_t)$  increases with the number of miners and decreases with the total transaction demand on the market. The probability of success for miners  $\lambda(\theta_t)$  increases with the transaction demand and decreases with the number of miners on the market.

$$\sigma(\theta_t) = \frac{T(d_t, CPU_t)}{d_t} = T\left(1, \frac{1}{\theta_t}\right) = \bar{T}\theta_t^{\eta-1}, \sigma'(\theta_t) < 0 \quad (4.20)$$

$$\lambda(\theta_t) = \frac{T(d_t, CPU_t)}{CPU_t} = T(\theta_t, 1) = \bar{T}\theta_t^\eta, \lambda'(\theta_t) > 0 \quad (4.21)$$

To sum up, too many transaction requests per miner implies an inefficiently small transaction processing probability in the model. In the real world, we observe it in an increase of transaction fees. Too many miners relative to the transaction demand implies an inefficiently low probability of a miner's success. In reality, higher competition on the mining market results in a higher energy consumption.

#### 4.3.2 Demand for Money and Transactions

Money demand arises from a double coincidence problem as in the Lagos and Wright (2005) search theoretic approach. I only briefly sketch the framework and refer the reader to the original paper or to Fernández-Villaverde and Sanches (2017) for the detailed derivations.

A continuum of buyers and a continuum of sellers, both of measure 1, live in the economy. Every period buyers are randomly assigned to sellers. Every buyer-seller pair negotiates on a deal according to which a seller produces and sells a particular amount of goods  $q_t$  to the buyer against a payment  $p_t$ . The buyer and the seller in every pair might never see each

---

<sup>40</sup>In case of the Bitcoin blockchain the difficulty is adjusted approximately once in 2 weeks.

other again. Consequently they cannot rely on a credit and have to use money as a means of payment to be able to trade.<sup>41</sup>

The negotiation of trade deals is achieved by "take-it-or-leave-it" offers from buyers to sellers. Denote the utility function of the buyer as  $u(q_t)$  and the disutility from work of the seller as  $w(n_t)$  where  $n_t$  is an amount of labor input.<sup>42</sup> The production function of the seller is linear and takes labor as the only input,  $q_t = n_t$ . Hence  $w(n_t) = w(q_t)$ . The first best amount of production and trade  $q^*$  is determined by  $u'(q^*) = w'(q^*)$ . If  $\beta$  is a discount factor,  $\phi_t$  is a value of a currency units in terms of real goods and  $m_t$  is money holdings of the buyer then the bargaining problem can be written as

$$\max_{q_t, p_t} [u(q_t) - \beta \phi_{t+1} p_t] \quad (4.22)$$

$$\text{s.t. } -w(q_t) + \beta \phi_{t+1} p_t \geq 0 \quad (4.23)$$

$$p_t \leq m_t \quad (4.24)$$

The buyer maximizes his utility minus the real value of the payment. The participation constraint (4.23) ensures that the seller wants to participate in the trade deal. It always holds with equality. The liquidity constraint (4.24) states that the payment cannot exceed the buyer's money holdings. Depending on whether the liquidity constraint is binding or not an interior or a corner solution is possible:

$$q_t = \begin{cases} q^* & \text{if } \phi_{t+1} m_t \geq \beta^{-1} w(q_t^*) \\ w^{-1}(\beta \phi_{t+1} m_t) & \text{if } \phi_{t+1} m_t < \beta^{-1} w(q_t^*) \end{cases}$$

$$\phi_{t+1} p_t = \begin{cases} \beta^{-1} w(q^*) & \text{if } \phi_{t+1} m_t \geq \beta^{-1} w(q_t^*) \\ \phi_{t+1} \hat{m}_t & \text{if } \phi_{t+1} m_t < \beta^{-1} w(q_t^*) \end{cases}$$

Whenever the buyer has an amount of real balances  $m^* = \beta^{-1} w(q^*)$  at hand he transfers this amount to the seller and consumes the optimal amount  $q^*$ . If the buyer has less real money balances at hand he simply spends all his money and gets whatever the seller is willing to produce for this payment.

In the standard Lagos-Wright model buyers and sellers are able to trade with an exogenous probability  $\sigma$ . In the current model trade also occurs only with the probability  $\sigma(\theta_t)$ .

---

<sup>41</sup>In the Lagos and Wright (2005) every agent produces a unique good. Money is required to overcome the double coincidence problem - that each party in the pair wants the good of the counter-party and the barter is possible. Exogenous distinction between buyers and sellers does not affect the results.

<sup>42</sup> $u(0) = 0, u'(0) = \infty, u'(\cdot) > 0, u''(\cdot) < 0$  and  $w(0) = 0, w'(\cdot) > 0, w''(\cdot) > 0$ .

However, the interpretation is different. In the current set-up the buyer-seller pair needs to transfer  $p_t$  currency units through the blockchain (one transaction is assumed to contain one currency unit). Each transaction is processed with the probability  $\sigma(\theta_t)$ . The pair thus submits  $d_t$  transaction requests such that  $p_t = \sigma(\theta_t)d_t$ . Since the buyer can only transfer currency units that he possesses, money demand equals the transaction demand,  $m_t = d_t$ .

As in the previous subsection denote as  $\gamma_t = \frac{\phi_{t+1}}{\phi_t}$  the return on money. Money demand stays finite only when  $\gamma_t \leq \beta^{-1}$ . Apart from the endogenous probability of trade, the money demand function is identical to Lagos, Wright (2005) and Fernández-Villaverde and Sanches (2017).

$$\sigma(\theta_t) \frac{u'(q_t)}{w'(q_t)} + 1 - \sigma(\theta_t) = \frac{1}{\beta\gamma_{t+1}} \quad (4.25)$$

$$\phi_{t+1}m_t = \beta^{-1}w(q_t) \quad (4.26)$$

Condition (4.26) comes from the bargaining solution. Condition (4.25) is defined by the utility maximization of a buyer. It determines the amount of trade as a function of the return on money given the trade probability. If  $\gamma_{t+1} = \beta^{-1}$  then  $q_t = q^*$ . This result corresponds to the Friedman (1969) rule.

### 4.3.3 Miners and Money Supply

The money market features free entry for miners. Every miner uses one computer processing unit CPU to process a transaction request and pays real costs  $P_t^e c \phi_t$ . Here  $P_t^e$  is the price of electricity in currency units and  $c$  is a fixed amount of electricity that one CPU consumes.

With probability  $\lambda(\theta_t)$  the miner "meets" a transaction request, processes it and receives  $r$  currency units as a reward. The reward can be spend in the next period and thus has a real value  $\phi_{t+1}r$ .

Free entry drives miners' profits to zero, therefore<sup>43</sup>

$$P_t^e c \phi_t = \beta \lambda(\theta_t) r \phi_{t+1} \quad (4.27)$$

If we were able to specify the electricity price in terms of consumer goods then we would be able to determine the real price of money. For example, if electricity is a consumption good itself then its price equals to the price level in the economy  $\frac{1}{\phi_t}$ . Then the (expected) future real price of money is driven by the electricity consumption of miners and their reward  $\phi_{t+1} = \frac{c}{\beta \lambda(\theta_t) r}$ .

---

<sup>43</sup>Miners are owned by buyers and sellers in equal proportions. Consequently, outside of the equilibrium path all profits and losses are equally redistributed among buyers and sellers.



Alternatively, the model needs to include an electricity market to describe the dynamics of  $P_t^e$ . I start with a simple version of the model by assuming that only a part of the total electricity demand comes from miners. Hence mining cannot significantly affect the electricity price. In other words I assume that the price is exogenous and constant:  $P_t^e = P^e \forall t$  and  $\psi \equiv P^e c$ . Note that nominal costs  $\psi$  and the nominal reward  $r$  are per transferred currency unit and  $0 < \psi < 1$  and  $0 < r < 1$ .

The free entry condition becomes:

$$\psi \phi_t = \beta \lambda(\theta_t) r \phi_{t+1} \quad (4.28)$$

Suppose that  $\phi_{t+1}$  increases and the profit of miners becomes positive. New miners enter the money market. Consequently, the market tightness declines and so does  $\lambda(\theta_t)$ . This drives the profits back to zero such that the condition (4.28) holds again.

Miners can not directly decide on the money supply. Money stock is increased when the processing of new transactions  $T_t$  is rewarded.

$$M_t = M_{t-1} + r T_t \quad (4.29)$$

Equation (4.29) can be written as  $M_t = M_{t-1} + r \sigma(\theta_t) d_t = M_{t-1} + r \sigma(\theta_t) m_t$ . On the equilibrium path money demand  $m_t$  and money supply  $M_t$  are equal.

$$M_t = M_{t-1} + r \sigma(\theta_t) M_t \quad (4.30)$$

$$M_t = \frac{1}{1 - r \sigma(\theta_t)} M_{t-1} \quad (4.31)$$

The model predicts a constant money growth in equilibrium. The growth rate is higher when 1) the reward per block  $r$  is higher or 2) the probability of the transaction processing  $\sigma(\theta_t)$  is higher. Since  $\sigma'(\theta_t) < 0$  the second condition means that the money growth rate is higher when the market tightness is lower.

**Proposition 1:** On the equilibrium trajectory money supply is growing with a positive rate. Constant money supply (and constant price) equilibrium is achieved only with a zero reward per block.

#### 4.3.4 Equilibrium Dynamics

Real money demand can be expressed as a function of the trade on the decentralized market  $q_t$  which is in turn a function of  $\gamma_{t+1}$  and  $\theta_t$ . From (4.26) one can express

$$z(\gamma_{t+1}, \theta_t) = m_t \phi_t = \frac{w(q_t(\gamma_{t+1}, \theta_t)) \phi_t}{\beta \phi_{t+1}} = \frac{w(q_t(\gamma_{t+1}, \theta_t))}{\beta \gamma_{t+1}} \quad (4.32)$$

where  $q_t$  is defined as (4.25). In equilibrium real money supply  $M_t\phi_t$  which is driven by (4.31) must be equal to the real money demand  $z(\gamma_{t+1}, \theta_t)$ . The dynamics of  $\theta_t$  follow (4.28). The equilibrium path is described by

$$z(\gamma_{t+1}, \theta_t) = \gamma_t z(\gamma_t, \theta_{t-1}) \frac{1}{1 - \sigma(\theta_t)r} \quad (4.33)$$

$$\psi = \beta r \gamma_{t+1} \lambda(\theta_t) \quad (4.34)$$

$$\psi = \beta r \gamma_t \lambda(\theta_{t-1}) \quad (4.35)$$

which defines  $\gamma_{t+1}(\gamma_t)$ . In contrast to Lagos and Wright (2005) demand for real money balances depends not only on the return on money but additionally on the money market tightness.

In contrast to the standard model, money demand does not go to 0 as  $\gamma_t$  goes to 0. From the equation (4.33) one can see that  $\gamma_t$  has two effects on the money demand  $z_{t+1}$ : a direct effect and the indirect effect via  $\theta_t$ . A decline in  $\gamma_t$  means that currency units have lower return. Consequently, the money demand declines. This purchasing power channel is the same as in the standard Lagos-Wright framework. On the other hand, lower  $\gamma_t$  means lower profit for the miners so the miners leave the money market. Market tightness rises and  $\sigma(\theta_t)$  falls. The probability of transaction processing decreases. Buyers, therefore, have to submit more transaction requests to be able to trade. Due to this precautionary motive buyers increase their money demand. The total result depends on the specification of functional forms and calibration.

To analyze the equilibrium explicitly I impose functional forms for the utility function and the function of the disutility from efforts. Following Fernández-Villaverde and Sanches, (2017) I use the following specifications:

$$u(q) = \frac{q^{1-g}}{1-g}, \text{ and } w(q) = \frac{q^{1+\alpha}}{1+\alpha}, \text{ where } 0 < g < 1, \alpha \geq 0. \quad (4.36)$$

The money demand function (4.37) resembles that of FVS (2016). However, the probability of trade  $\sigma(\theta_t)$  is endogenous and determined by (4.20).  $\theta_t$  can be expressed as a function of  $\gamma_{t+1}$  from (4.34)

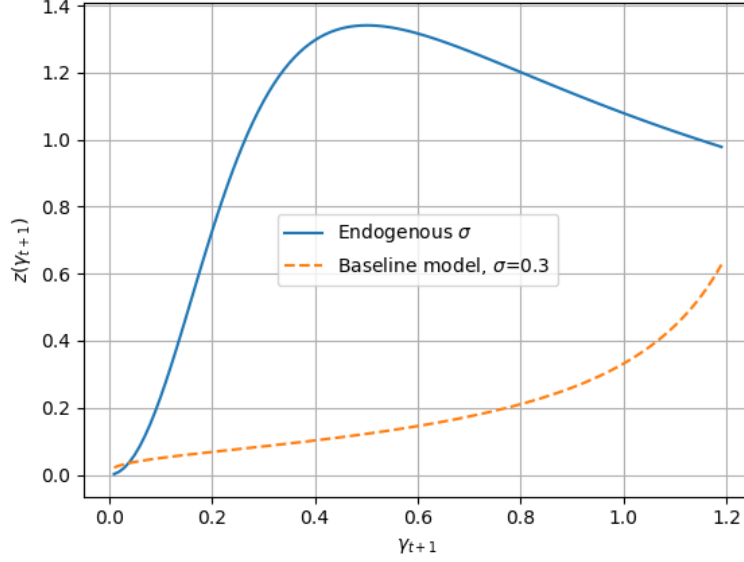
$$z(\gamma_{t+1}, \theta_t) = \frac{(\beta \gamma_{t+1})^{\frac{1+\alpha}{g+\alpha}-1}}{1+\alpha} \left[ \frac{\sigma(\theta_t)}{1 - (1 - \sigma(\theta_t))\beta \gamma_{t+1}} \right]^{\frac{1+\alpha}{g+\alpha}} \quad (4.37)$$

Figure (4.1) plots money demand (4.37). I took the standard parameter values:  $\beta = 0.997$ ,  $\alpha = 0.5$ ,  $\eta = 0.5$ ,  $g = 0.5$  I chose the level of reward based on the statistics for the BTC blockchain. According to blockchain.info<sup>44</sup> a miner's revenue lies between 0.6% and 1.8% of

---

<sup>44</sup><https://blockchain.info/charts/cost-per-transaction-percent>

a transaction volume. I set  $r=1.5\%$ . Mining costs per block are 20% which means  $\psi = 0.2$ . While it is difficult to estimate the electricity amount costs of miners the resulting money demand function is pretty insensitive to the value of  $\psi$ .



**Figure 4.1:** Money Demand as a Function of the Return on Money.

The dashed line presents the demand function for a constant probability of transaction confirmation  $\sigma = 0.3$ . It corresponds to the money demand function in the baseline model of Fernández-Villaverde and Sanches (2016). In this case money demand uniformly increases with  $\gamma_{t+1}$ . This result comes from the purchasing power channel. The solid line plots the money demand in the current model. It shows that after  $\gamma_{t+1}$  reaches a point around 0.5, the demand for money decreases. The reason for that is the presence of mining. When deciding on the money holdings and the transaction demand buyers take into account that only  $\sigma_t d_t$  of currency units will be transferred. Because  $\sigma(\theta_t)$  is less than one, buyers hold an excessive amount of currency and submit an excessive number of transactions. This can be thought of as a precautionary money demand channel.

The gap between two functions quantitatively depends on the chosen parameter values. However, qualitatively the result is robust to different calibrations. The elasticity of substitution in the matching function most strongly affects the form of  $z(\gamma_{t+1})$ . The humped shape is preserved for non-extreme values,  $\eta \in [0.3, 0.8]$ .

The additional precautionary channel in the money demand changes the stability properties of the monetary equilibrium if  $\gamma$  in this equilibrium corresponds to the declining part of

$z(\gamma_{t+1})$ . To see that, substitute the specifications (4.36) into the equilibrium conditions (4.33) - (4.35) to obtain the equilibrium trajectory  $\gamma_{t+1}(\gamma_t)$ :

$$\frac{\gamma_{t+1}^{\frac{1+\alpha}{g+\alpha} \frac{1}{\eta}-1} \left[ 1 - r\bar{T} \left( \frac{\psi}{T\beta\gamma_{t+1}} \right)^{\frac{\eta-1}{\eta}} \right]}{\left[ 1 - \left( 1 - T \left( \frac{\psi}{T\beta\gamma_{t+1}} \right)^{\frac{\eta-1}{\eta}} \right) \beta\gamma_{t+1} \right]^{\frac{1+\alpha}{g+\alpha}}} = \frac{\gamma_t^{\frac{1+\alpha}{g+\alpha} \frac{1}{\eta}}}{\left[ 1 - \left( 1 - T \left( \frac{\psi}{T\beta\gamma_t} \right)^{\frac{\eta-1}{\eta}} \right) \beta\gamma_t \right]^{\frac{1+\alpha}{g+\alpha}}} \quad (4.38)$$

This expression describes the dynamics of the return on money  $\gamma_{t+1}$  as a function of  $\gamma_t$ . First, note that  $\gamma=0$  is a solution of this equation. There exists a real or non-monetary equilibrium with zero return on money, zero money demand and no trade in the decentralized market.

In addition to the real equilibrium, there exists a monetary equilibrium with  $\gamma > 0$  such that

$$A\gamma^{\frac{1-\eta}{\eta}} + \gamma - 1 = 0 \quad (4.39)$$

$$\text{where } A = (Tr)^{\frac{1}{\eta}} \beta^{\frac{1-\eta}{\eta}} \psi^{\frac{\eta-1}{\eta}} > 0 \quad (4.40)$$

Let us consider a special case with  $\eta=0.5$ . In this case the model can be solved analytically and the equilibrium return on money,  $\gamma^*$  is defined as:

$$\gamma^* = \frac{1}{1 + \frac{\beta T^2 r^2}{\psi}} = \frac{\psi}{\psi + \beta T^2 r^2} \quad (4.41)$$

It follows from the (4.41) that the value of  $\gamma^*$  depends on the ratio  $\frac{r^2}{\psi}$ .

**Proposition 2 (comparative statics):** Under functional forms (4.36) the return on money in a monetary equilibrium 1) increases in the cost parameter  $\psi$ , 2) decreases in the reward  $r$ .

Higher mining costs discourage miners from processing transactions and hence a higher return on money  $\gamma^*$  is required to compensate the miners. In the extreme case if  $\psi$  goes to zero and  $r > 0$  the miners receive a positive reward at zero costs. Infinitely many miners enter the money market and the probability of success of a single miner goes to zero. The number of mined blocks goes to zero and the currency becomes valueless,  $\gamma^* = 0$ .

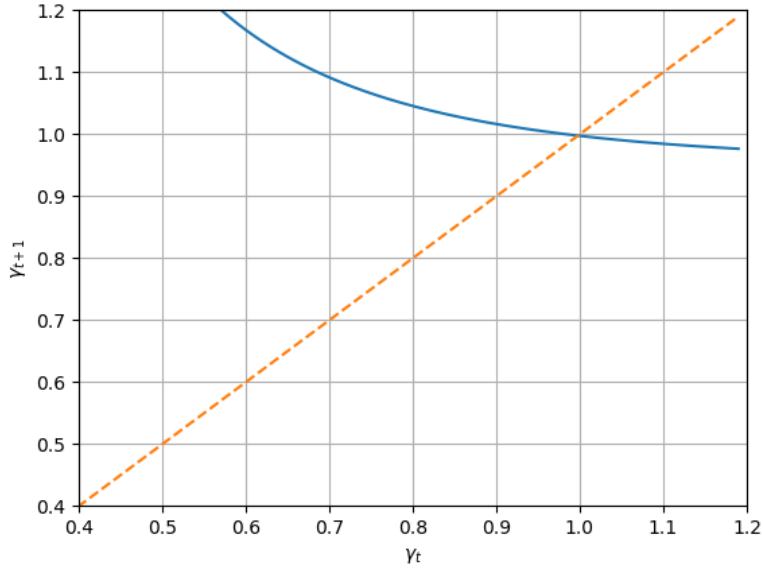
Higher reward  $r$  incentivizes more miners to enter the market. More currency units are injected into the economy which leads to a lower return on money. If the reward to miners goes to zero, then there is no injections of newly created currency units and the return on money stays constant. It means that the price level stays constant.

**Proposition 3:** In a monetary equilibrium with finite mining costs the return on money  $\gamma^*$  is less than 1 which means that inflation is positive as long as the reward to miners is positive.

This is an important feature of the blockchain cryptocurrencies that the model is able to correctly reproduce.

A blockchain-based monetary system is unable to achieve price stability. The presence of electricity costs requires a reward to miners in a form of newly created currency units. Consequently, the money supply is rising by construction. Moreover, if the blockchain protocol specifies that reward will go to zero at some point of time and miners will be compensated solely by the transaction fees, then  $\gamma^*$  will converge to one and a positive money growth rate will no longer be necessary.

Let us consider equation (4.38) again which describes equilibria as  $\gamma_{t+1} = \gamma_t = \gamma^*$ . Figure (4.2) plots the function  $\gamma_{t+1}(\gamma_t)$  for the same calibration as described above together with a 45-degree line. The equilibrium return on money is given by points of interception of two lines. For the calibration employed,  $\gamma^* = 0.997$  which corresponds to approximately 3% inflation.



**Figure 4.2:** Monetary Steady State

The slope of the function at  $\gamma^*$  is smaller than 1. It means that the equilibrium trajectories that start from  $\gamma_0 > 0$  converge to monetary equilibrium with a positive value of currency. Because  $\gamma^*$  is on the declining segment of the money demand function the precautionary channel plays a major role. The return on money declines when the initial value is too high and rises when the initial value is too low. As before this result holds for  $\eta \in [0.3, 0.8]$ .

**Proposition 4:** The monetary equilibrium is stable due to the presence of a precautionary money demand channel.

## 4.4 Conclusion

Monetary economics has studied currency competition since at least free banking era episodes in the USA and Scotland. Recent development of cryptocurrencies has revived the discussion on the outcome of currency competition.

Cryptocurrencies are "controlled" by private rational agents in accordance with a specific protocol - the blockchain. Due to distributed trust concerns the functioning of the blockchain requires a significant input of a computational time and, thus, energy. Moreover, operation of the blockchain features externalities.

This chapter develops two monetary models that can be applied to the analysis of digital currencies. The first part of the chapter studies how the outcome of currency competition is affected by the costs associated with private currency circulation. I extend the model of Fernandez-Villaverde and Sanches (2016) with linear operational costs. The demand side of the money market is based on the Lagos and Wright (2005). I analyze equilibrium with a purely private money provision and with a competition between private and public currencies.

The equilibrium with private currency provision does not feature price stability. A positive growth rate for private money supply and a positive seigniorage is needed to compensate private currency providers for their activities. The costs of private money circulation lead to an inefficient level of production and trade on the decentralized market. The magnitude of the reduction in production and trade is proportional to the cost parameter.

Even if currency competition does not deliver price stability it might impose discipline on the public money. I show that currency competition plays a disciplinary role for the return on public money (and, thus, on inflation). However, this is true only if the operational costs of private cryptocurrencies are below a certain level. The role of currency competition, therefore, depends on the technology that underlies a circulation of private currencies.

The second part of the essay develops a monetary model of a private digital currency circulating according to the blockchain protocol. I propose to use a matching function to model how miners "meet" pending transaction requests and "create" new processed transactions on the blockchain. The matching function naturally captures the externalities of the blockchain monetary system: the probability of a miner's success increases with the transaction demand and decreases with the total number of miners on the market; the

transaction processing probability increases with the number of miners and declines with the number of pending transaction requests.

The standard framework is altered by the fact that agents have to submit transaction requests before they are able to transfer currency units. Every transaction request is processed only with a particular probability and hence the trade also happens only with a particular probability. In contrast to the Lagos-Wright approach in the current set-up the probability of trade is endogenous. It depends on the number of miners and the transaction demand.

I use the model to analyze the effects of mining costs and rewards on the equilibrium outcome. The return on money in equilibrium increases with the mining costs and decreases with the reward for miners. Higher mining costs require a higher return on currency to compensate miners for their work and a higher reward means that a lower return on money is sufficient.

Endogenous probability of transaction processing and, therefore, of trade changes the properties of the money demand function. More specifically, agents hold an excessive amount of money units due to a precautionary motive. The reason for the precautionary money demand is the fact that every transaction is processed only with a particular probability. With an increase in the return on money more miners start to operate and the probability of transaction processing rises. Agents reduce their excessive money holdings and money demand goes down. In contrast to the standard model in which money demand uniformly increases with the return on money, in the current set up money demand is hump-shaped: it increases for low values of the return on money and starts to decline after a particular level when the precautionary motive prevails.

The hump-shaped money demand function changes the stability properties of the system. If in the monetary equilibrium the return on money corresponds to the declining part of the money demand function then this equilibrium is stable. In other words, money cannot become valueless if its initial value is positive. If, for example, a shock forces the return on money to go below the equilibrium level then some miners leave the market. Individuals then have to increase their precautionary money demand which puts an upward pressure on the money return. Similarly, the return on money cannot explode.

The proposed models abstract away many important and interesting ingredients of actual blockchain-based currencies. For instance, decisions to attach transaction fees to transaction requests, miners' decisions about the composition of the block, costs of entering the mining market, adjustments in the difficulty or a necessity to possess a particular amount of currency

units to be allowed to mine. In reality miners do not directly decide on the money supply but on the mining intensity instead. In fact, for many cryptocurrencies the evolution of the money supply is written directly in the code of the blockchain protocol. From the economic point of view the paper abstracts from the social costs of the excessive electricity consumption, the capital overinvestment and the environmental pollution. Additional features can be introduced into the current framework at the costs of reducing the model tractability. Many of these aspects are novel to the literature and their accurate formulation represents an independent research topic on its own.



## 4.5 Appendix

### Buyer Objective Function

The first derivative of the buyer objective function (for any currency  $i$ ) can be written as

$$-\phi_t + \beta\phi_{t+1} + \sigma[u'(q) - w'(q)]\frac{1}{w'(q)}$$

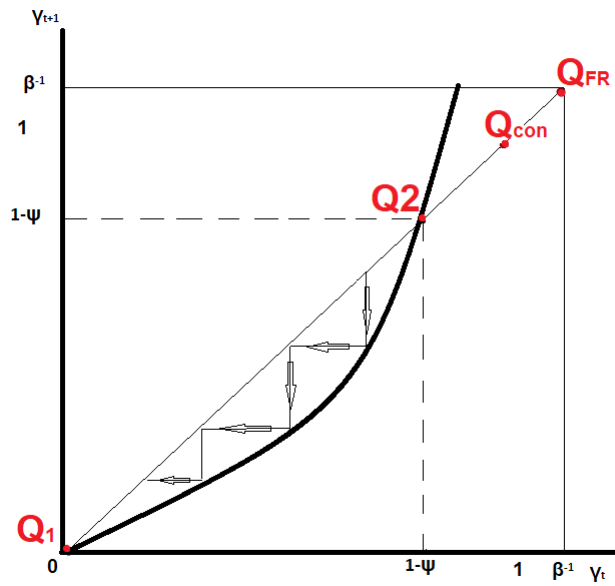
As  $m$  approaches  $\beta^{-1}w(q^*)$  from below,  $q(m_t)$  goes to  $q^*$  and the third term goes to zero. The first term is negative for  $\frac{\phi_{t+1}}{\phi_t} < \beta^{-1}$ . Therefore, the derivative is negative and the objective function is decreasing at  $\beta^{-1}w(q^*)$ .

The second derivative of the objective function can be written as

$$\sigma u''[q(m_t)](q'(m_t))^2 + \sigma u'[q(m_t)]q''(m_t)$$

Using an implicit function theorem one can show that the second derivative is negative and the objective function is concave since  $u''(q) < 0$ ,  $u'(q) > 0$ ,  $q'(m_t) = \frac{1}{w'(q)} > 0$  and  $q''(m_t) = -\frac{w''(q)}{w'(q)^3} < 0$ . The buyer problem has a unique solution at  $m_t < \beta^{-1}w(q^*)$ .

## Plot of the Equilibrium Dynamics



**Figure 4.3:** Plot of the Equilibrium Dynamics

*Note:* The 45-degree line represents the saddle points of the system.  $\gamma_{t+1}(\gamma_t)$  plots the transition equation of the economy.  $Q_1$  stands for cashless equilibrium and  $Q_2$  is an unstable monetary equilibrium,  $Q_{con}$  is a constant price equilibrium and  $Q_{FR}$  denotes the Friedman rule equilibrium.

## Threshold Values for Cost Parameter

<b>Table 4.1:</b> Threshold Values for Cost Parameter											
$\eta$	0.1	0.1	0.1	0.1	0.1	0.1	0.3	0.3	0.3	0.5	0.5
$\alpha$	0.1	0.1	0.1	0.3	0.3	0.3	0.1	0.1	0.3	0.5	0.3
$\sigma$	0.1	0.3	0.5	0.1	0.3	0.5	0.1	0.3	0.1	0.1	0.1
$\gamma^{G*}$	0.97	0.92	0.86	0.95	0.85	0.8	0.94	0.8	0.91	0.88	0.84
$\psi^{max}$	0.03	0.08	0.14	0.05	0.15	0.2	0.06	0.2	0.09	0.12	0.16





## References

- AHMED, N. K., A. F. ATIYA, N. E. GAYAR, AND H. EL-SHISHINY (2010): “An empirical comparison of machine learning models for time series forecasting,” *Econometric Reviews*, 29, 594–621.
- ALCIDÌ, C., A. FLAMINI, AND A. FRACASSO (2011): “Policy regime changes, judgment and Taylor rules in the Greenspan era,” *Economica*, 78, 89–107.
- ALVAREZ-LOIS, P. P. (2000): “Asymmetries in the capacity-inflation trade-off,” .
- AMISANO, G. AND O. TRISTANI (2010): “Euro area inflation persistence in an estimated nonlinear DSGE model,” *Journal of Economic Dynamics and Control*, 34, 1837–1858.
- APARICIO, D. AND M. BERTOLOTTO (2017): “Forecasting inflation with online prices,” .
- ARRAS, L., G. MONTAVON, K.-R. MÜLLER, AND W. SAMEK (2017): “Explaining recurrent neural network predictions in sentiment analysis,” *arXiv preprint arXiv:1706.07206*.
- ARUOBA, S. B., L. BOCOLA, AND F. SCHORFHEIDE (2017): “Assessing DSGE model nonlinearities,” *Journal of Economic Dynamics and Control*, 83, 34–54.
- ATKESON, A. AND L. E. OHANIAN (2001): “Are Phillips curves useful for forecasting inflation?” *Quarterly Review*, 2–11.
- BARNICHON, R. AND C. MATTHES (2016): “Gaussian mixture approximations of impulse responses and the non-linear effects of monetary shocks,” .
- BARRON, A. R. (1993): “Universal approximation bounds for superpositions of a sigmoidal function,” *IEEE Transactions on Information theory*, 39, 930–945.
- BERENTSEN, A. (2006): “On the private provision of fiat currency,” *European Economic Review*, 50, 1683–1698.
- BRÜGGEMANN, R. AND J. RIEDEL (2011): “Nonlinear interest rate reaction functions for the UK,” *Economic Modelling*, 28, 1174–1185.
- BUNZEL, H. AND W. ENDERS (2010): “The Taylor rule and “opportunistic” monetary policy,” *Journal of Money, Credit and Banking*, 42, 931–949.

- CAVALCANTI, R. D. O., A. EROSA, AND T. TEMZELIDES (1999): “Private money and reserve management in a random-matching model,” *Journal of Political Economy*, 107, 929–945.
- CAVALCANTI, R. D. O. AND N. WALLACE (1999): “A model of private bank-note issue,” *Review of Economic Dynamics*, 2, 104–136.
- CHEN, X., J. RACINE, AND N. R. SWANSON (2001): “Semiparametric ARX neural-network models with an application to forecasting inflation,” *IEEE Transactions on neural networks*, 12, 674–683.
- CUKIERMAN, A. AND S. GERLACH (2003): “The inflation bias revisited: theory and some international evidence,” *Working Paper*.
- CUKIERMAN, A. AND A. MUSCATELLI (2002): “Do central banks have precautionary demands for expansions and for price stability?-Theory and evidence,” *Working Paper*.
- CYBENKO, G. (1989): “Approximation by superpositions of a sigmoidal function,” *Mathematics of control, signals and systems*, 2, 303–314.
- DEL NEGRO, M. AND F. SCHORFHEIDE (2004): “Priors from general equilibrium models for VARs,” *International Economic Review*, 45, 643–673.
- DIEBOLD, F. X. AND R. S. MARIANO (2002): “Comparing predictive accuracy,” *Journal of Business & economic statistics*, 20, 134–144.
- DIJK, D. V., T. TERÄSVIRTA, AND P. H. FRANSES (2002): “Smooth transition autoregressive models—a survey of recent developments,” *Econometric reviews*, 21, 1–47.
- DOLADO, J., R. M.-D. PEDRERO, AND F. J. RUGE-MURCIA (2004): “Nonlinear monetary policy rules: some new evidence for the US,” *Studies in Nonlinear Dynamics & Econometrics*, 8.
- DOLADO, J. J., R. MARIA-DOLORES, AND M. NAVEIRA (2005): “Are monetary-policy reaction functions asymmetric?: The role of nonlinearity in the Phillips curve,” *European Economic Review*, 49, 485–503.
- DUBITZKY, W., M. GRANZOW, AND D. P. BERRAR (2007): *Fundamentals of data mining in genomics and proteomics*, Springer Science & Business Media.

- ELGER, T., J. BINNER, B. NILSSON, AND J. TEPPER (2006): “Predictable non-linearities in U.S. Inflation,” *Economics Letters*, 93, 323–328.
- FERNÁNDEZ-VILLAYERDE, J. AND J. F. RUBIO-RAMÍREZ (2007): “Estimating macroeconomic models: A likelihood approach,” *The Review of Economic Studies*, 74, 1059–1087.
- FERNÁNDEZ-VILLAYERDE, J. AND J. F. RUBIO-RUBIO-RAMÍREZ (2004): “Comparing dynamic equilibrium models to data: a Bayesian approach,” *Journal of Econometrics*, 123, 153–187.
- FERNÁNDEZ-VILLAYERDE, J. AND D. SANCHES (2016): “Can currency competition work?” Tech. rep., National Bureau of Economic Research.
- FINOCCHIARO, D. AND V. Q. VON HEIDEKEN (2013): “Do central banks react to house prices?” *Journal of Money, Credit and Banking*, 45, 1659–1683.
- FRIEDMAN, M. (1969): “The optimum quantity of money and other essays,” .
- GERLACH, S. AND J. LEWIS (2014): “Zero lower bound, ECB interest rate policy and the financial crisis,” *Empirical Economics*, 46, 865–886.
- GEWEKE, J. (1999): “Using simulation methods for Bayesian econometric models: inference, development, and communication,” *Econometric reviews*, 18, 1–73.
- HARVEY, D., S. LEYBOURNE, AND P. NEWBOLD (1997): “Testing the equality of prediction mean squared errors,” *International Journal of forecasting*, 13, 281–291.
- HAYEK, F. A. (1976): *Denationalization of Money—The Argument Refined: An Analysis of the Theory and Practice of Concurrent Currencies*, Institute for Economic Affairs.
- HOCHREITER, S. AND J. SCHMIDHUBER (1997): “Long short-term memory,” *Neural computation*, 9, 1735–1780.
- JACK, W. AND T. SURI (2011): “Mobile money: The economics of M-PESA,” Tech. rep., National Bureau of Economic Research.
- KESRIYELI, M., D. OSBORN, AND M. SENSIER (2004): “Nonlinear and structure change in interest reaction function for the US, UK and Germany,” *Discussion Paper Series of the University of Manchester*.



- KIM, J., S. KIM, E. SCHAUMBURG, AND C. A. SIMS (2008): “Calculating and using second-order accurate solutions of discrete time dynamic equilibrium models,” *Journal of Economic Dynamics and Control*, 32, 3397–3414.
- KIM, J. AND F. J. RUGE-MURCIA (2009): “How much inflation is necessary to grease the wheels?” *Journal of Monetary Economics*, 56, 365–377.
- KINGMA, D. P. AND J. BA (2014): “Adam: A Method for Stochastic Optimization,” *CoRR*, abs/1412.6980.
- KLEIN, B. (1974): “The competitive supply of money,” *Journal of Money, Credit and Banking*, 6, 423–453.
- KOCK, A. B., T. TERÄSVIRTA, ET AL. (2011): “Forecasting macroeconomic variables using neural network models and three automated model selection techniques,” Tech. rep., Department of Economics and Business Economics, Aarhus University.
- KOOP, G., M. H. PESARAN, AND S. M. POTTER (1996): “Impulse response analysis in nonlinear multivariate models,” *Journal of econometrics*, 74, 119–147.
- KUAN, C.-M. AND T. LIU (1995): “Forecasting exchange rates using feedforward and recurrent neural networks,” *Journal of applied econometrics*, 10, 347–364.
- LAGOS, R. AND R. WRIGHT (2005): “A unified framework for monetary theory and policy analysis,” *Journal of political Economy*, 113, 463–484.
- LAPUSCHKIN, S., A. BINDER, G. MONTAVON, K.-R. MÜLLER, AND W. SAMEK (2016): “The LRP toolbox for artificial neural networks,” *The Journal of Machine Learning Research*, 17, 3938–3942.
- LAXTON, D., G. MEREDITH, AND D. ROSE (1995): “Asymmetric effects of economic activity on inflation: Evidence and policy implications,” *Staff papers*, 42, 344–374.
- LAXTON, D., D. ROSE, AND D. TAMBAKIS (1999): “The US Phillips curve: The case for asymmetry,” *Journal of Economic dynamics and Control*, 23, 1459–1485.
- LO, M. C. AND J. PIGER (2005): “Is the response of output to monetary policy asymmetric? Evidence from a regime-switching coefficients model,” *Journal of Money, Credit and Banking*, 865–886.

- LOMBARDO, G. (2010): “On approximating DSGE models by series expansions,” *Working paper*.
- LUBIK, T. A. AND F. SCHORFHEIDE (2007): “Do central banks respond to exchange rate movements? A structural investigation,” *Journal of Monetary Economics*, 54, 1069–1087.
- MARCELLINO, M., J. H. STOCK, AND M. W. WATSON (2006): “A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series,” *Journal of econometrics*, 135, 499–526.
- MARIMON, R., J. P. NICOLINI, AND P. TELES (2003): “Inside–outside money competition,” *Journal of Monetary Economics*, 50, 1701–1718.
- (2012): “Money is an experience good: Competition and trust in the private provision of money,” *Journal of Monetary Economics*, 59, 815–825.
- MARKOV, N. AND C. DE PORRES (2012): “Is the Taylor Rule nonlinear? Empirical evidence from a semi-parametric modeling approach,” .
- MARTIN, C. AND C. MILAS (2004): “Modelling monetary policy: inflation targeting in practice,” *Economica*, 71, 209–221.
- MCADAM, P. AND P. MCNELIS (2005): “Forecasting inflation with thick models and neural networks,” *Economic Modelling*, 22, 848–867.
- MEYER, L. H., E. T. SWANSON, AND V. W. WIELAND (2001): “NAIRU uncertainty and nonlinear policy rules,” *American Economic Review*, 91, 226–231.
- NAIR, V. AND G. E. HINTON (2010): “Rectified linear units improve restricted boltzmann machines,” in *Proceedings of the 27th international conference on machine learning (ICML-10)*, 807–814.
- NAKAMURA, E. (2005): “Inflation forecasting using a neural network,” *Economics Letters*, 86, 373–378.
- O'DWYER, K. J. AND D. MALONE (2014): “Bitcoin mining and its energy footprint,” .
- ORPHANIDES, A. AND V. WIELAND (2000): “Efficient monetary policy design near price stability,” *Journal of the Japanese and International Economies*, 14, 327–365.
- PISSARIDES, C. A. (2000): *Equilibrium unemployment theory*, MIT press.

- QIN, T. AND W. ENDERS (2008): “In-sample and out-of-sample properties of linear and nonlinear Taylor rules,” *Journal of Macroeconomics*, 30, 428–443.
- ROTEMBERG, J. J. (1982): “Sticky prices in the United States,” *Journal of Political Economy*, 90, 1187–1211.
- RUGE-MURCIA, F. (2014): “Indirect inference estimation of nonlinear dynamic general equilibrium models: with an application to asset pricing under skewness risk,” Tech. rep., Centre interuniversitaire de recherche en économie quantitative, CIREQ.
- RUGE-MURCIA, F. J. (2004): “The inflation bias when the central bank targets the natural rate of unemployment,” *European Economic Review*, 48, 91–107.
- RUGE-MURCIA, F. J. AND L. DUDLEY (2001): *Inflation targeting under asymmetric preferences*, International Monetary Fund, Research Department.
- SARGENT, T. J. (2009): *Dynamic macroeconomic theory*, Harvard University Press.
- SCHMITT-GROHÉ, S. AND M. URIBE (2004): “Solving dynamic general equilibrium models using a second-order approximation to the policy function,” *Journal of economic dynamics and control*, 28, 755–775.
- SCHORFHEIDE, F. (2000): “Loss function-based evaluation of DSGE models,” *Journal of Applied Econometrics*, 15, 645–670.
- SIMS, C. A. AND T. ZHA (2006): “Were there regime switches in US monetary policy?” *American Economic Review*, 96, 54–81.
- SMETS, F. AND R. WOUTERS (2003): “An estimated dynamic stochastic general equilibrium model of the euro area,” *Journal of the European economic association*, 1, 1123–1175.
- (2007): “Shocks and frictions in US business cycles: A Bayesian DSGE approach,” *American economic review*, 97, 586–606.
- STOCK, J. H. AND M. W. WATSON (1998): “A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series,” Working Paper 6607, National Bureau of Economic Research.
- (1999): “Forecasting inflation,” *Journal of Monetary Economics*, 44, 293–335.
- (2007): “Why has US inflation become harder to forecast?” *Journal of Money, Credit and banking*, 39, 3–33.

- SUNDERMEYER, M., R. SCHLÜTER, AND H. NEY (2012): “LSTM neural networks for language modeling,” in *Thirteenth annual conference of the international speech communication association*.
- SURICO, P. (2002): *Asymmetric central bank preferences and nonlinear policy rules*, Università Bocconi.
- SWANSON, N. R. AND H. WHITE (1997a): “Forecasting economic time series using flexible versus fixed specification and linear versus nonlinear econometric models,” *International journal of Forecasting*, 13, 439–461.
- (1997b): “A model selection approach to real-time macroeconomic forecasting using linear models and artificial neural networks,” *Review of Economics and Statistics*, 79, 540–550.
- TAUB, B. (1985): “Private fiat money with many suppliers,” *Journal of Monetary Economics*, 16, 195–208.
- TAYLOR, J. B. (1993): “Discretion versus policy rules in practice,” in *Carnegie-Rochester conference series on public policy*, Elsevier, vol. 39, 195–214.
- TERÄSVIRTA, T. (2006): “Forecasting economic variables with nonlinear models,” *Handbook of economic forecasting*, 1, 413–457.
- TIKHONOV, A. N. (1963): “On the solution of ill-posed problems and the method of regularization,” in *Doklady Akademii Nauk*, Russian Academy of Sciences, vol. 151, 501–504.
- VARIAN, H. R. (1975): “A Bayesian approach to real estate assessment,” *Studies in Bayesian Econometric and Statistics in honor of Leonard J. Savage*, 195–208.



## Selbstständigkeitserklärung

Ich versichere, die von mir vorgelegte Dissertation selbstständig und ohne unerlaubte Hilfe und Hilfsmittel angefertigt, sowie die benutzten Quellen und Daten anderen Ursprungs als solche kenntlich gemacht zu haben.

Ich bezeuge durch meine Unterschrift, dass meine Angaben über die bei der Abfassung meiner Dissertation benutzten Hilfsmittel, über die mir zuteil gewordene Hilfe sowie über frühere Begutachtungen meiner Dissertation in jeder Hinsicht der Wahrheit entsprechen.

Berlin, 30. November 2018

Anna Almosova